

Am

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
12 July 2001 (12.07.2001)

PCT

(10) International Publication Number
WO 01/49715 A2(51) International Patent Classification?: **C07K 14/00**

(21) International Application Number: PCT/US00/30952

(22) International Filing Date:
8 November 2000 (08.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

PCT/US00/00376	6 January 2000 (06.01.2000)	US
PCT/US00/04342		
	18 February 2000 (18.02.2000)	US
PCT/US00/05841	2 March 2000 (02.03.2000)	US
PCT/US00/08439	30 March 2000 (30.03.2000)	US
PCT/US00/20710	28 July 2000 (28.07.2000)	US

(71) Applicant (for all designated States except US): **GENENTECH, INC.** [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ASHKENAZI, Avi, J.** [US/US]; 1456 Tarrytown Street, San Mateo, CA 94402 (US). **GODDARD, Audrey** [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). **GURNEY, Austin, L.** [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). **NAPIER, Mary, A.** [US/US]; 1015 Hayne Road,Hillsborough, CA 94010 (US). **WATANABE, Colin, K.** [US/US]; 128 Corliss Drive, Moraga, CA 94556 (US). **WOOD, William, I.** [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US).(74) Agents: **BARNES, Elizabeth M.** et al.; c/o Genentech, Inc., MS49, 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH

(57) Abstract: The present invention concerns methods and compositions for inhibiting neoplastic cell growth. In particular, the present invention concerns antitumor compositions and methods for the treatment of tumors. The invention further concerns screening methods for identifying growth inhibitory, e.g., antitumor compounds. The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

WO 01/49715 A2

METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH

FIELD OF THE INVENTION

The present invention concerns methods and compositions for inhibiting neoplastic cell growth. In particular, the present invention concerns antitumor compositions and methods for the treatment of tumors. The invention further concerns screening methods for identifying growth inhibitory, *e.g.*, antitumor compounds.

BACKGROUND OF THE INVENTION

Malignant tumors (cancers) are the second leading cause of death in the United States, after heart disease (Boring *et al.*, CA Cancer J. Clin., 43:7 (1993)).

Cancer is characterized by the increase in the number of abnormal, or neoplastic, cells derived from a normal tissue which proliferate to form a tumor mass, the invasion of adjacent tissues by these neoplastic tumor cells, and the generation of malignant cells which eventually spread via the blood or lymphatic system to regional lymph nodes and to distant sites (metastasis). In a cancerous state a cell proliferates under conditions in which normal cells would not grow. Cancer manifests itself in a wide variety of forms, characterized by different degrees of invasiveness and aggressiveness.

Despite recent advances in cancer therapy, there is a great need for new therapeutic agents capable of inhibiting neoplastic cell growth. Accordingly, it is the objective of the present invention to identify compounds capable of inhibiting the growth of neoplastic cells, such as cancer cells.

SUMMARY OF THE INVENTION

A. Embodiments

The present invention relates to methods and compositions for inhibiting neoplastic cell growth. More particularly, the invention concerns methods and compositions for the treatment of tumors, including cancers, such as breast, prostate, colon, lung, ovarian, renal and CNS cancers, leukemia, melanoma, etc., in mammalian patients, preferably humans.

In one aspect, the present invention concerns compositions of matter useful for the inhibition of neoplastic cell growth comprising an effective amount of a PRO polypeptide as herein defined, or an agonist thereof, in admixture with a pharmaceutically acceptable carrier. In a preferred embodiment, the composition of matter comprises a growth inhibitory amount of a PRO polypeptide, or an agonist thereof. In another preferred embodiment, the composition comprises a cytotoxic amount of a PRO polypeptide, or an agonist thereof. Optionally, the compositions of matter may contain one or more additional growth inhibitory and/or cytotoxic

and/or other chemotherapeutic agents.

In a further aspect, the present invention concerns compositions of matter useful for the treatment of a tumor in a mammal comprising a therapeutically effective amount of a PRO polypeptide as herein defined, or an agonist thereof. The tumor is preferably a cancer.

5 In another aspect, the invention concerns a method for inhibiting the growth of a tumor cell comprising exposing the cell to an effective amount of a PRO polypeptide as herein defined, or an agonist thereof. In a particular embodiment, the agonist is an anti-PRO agonist antibody. In another embodiment, the agonist is a small molecule that mimics the biological activity of a PRO polypeptide. The method may be performed *in vitro* or *in vivo*.

10 In a still further embodiment, the invention concerns an article of manufacture comprising:

(a) a container;

(b) a composition comprising an active agent contained within the container; wherein the composition is effective for inhibiting the neoplastic cell growth, *e.g.*, growth of tumor cells, and the active agent in the composition is a PRO polypeptide as herein defined, or an agonist thereof; and

15 (c) a label affixed to said container, or a package insert included in said container referring to the use of said PRO polypeptide or agonist thereof, for the inhibition of neoplastic cell growth, wherein the agonist may be an antibody which binds to the PRO polypeptide.

In a particular embodiment, the agonist is an anti-PRO agonist antibody. In another embodiment, the agonist is a small molecule that mimics the biological activity of a PRO polypeptide. Similar articles of manufacture comprising a PRO polypeptide as herein defined, or an agonist thereof, in an amount that is therapeutically effective for the treatment of tumor are also within the scope of the present invention. Also within the scope of the invention are articles of manufacture comprising a PRO polypeptide as herein defined, or an agonist thereof, and a further growth inhibitory agent, cytotoxic agent or chemotherapeutic agent.

25 B. Additional Embodiments

In other embodiments of the present invention, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a PRO polypeptide.

In one aspect, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity,

alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule encoding a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule comprising the coding sequence of a full-length PRO polypeptide cDNA as disclosed herein, the coding sequence of a PRO polypeptide lacking the signal peptide as disclosed herein, the coding sequence of an extracellular domain of a transmembrane PRO polypeptide, with or without the signal peptide, as disclosed herein or the coding sequence of any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In a further aspect, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule that encodes the same mature polypeptide encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect of the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptides are disclosed herein. Therefore, soluble extracellular domains of the herein described PRO polypeptides are contemplated.

Another embodiment is directed to fragments of a PRO polypeptide coding sequence, or the complement thereof, that may find use as, for example, hybridization probes, for encoding fragments of a PRO polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-PRO antibody or as antisense oligonucleotide probes. Such nucleic acid fragments are usually at least about 20 nucleotides in length, alternatively at least about 30 nucleotides in length, alternatively at least about 40 nucleotides in length, alternatively at least about 50 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 70 nucleotides in length, alternatively at least about 80 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 100 nucleotides in length, alternatively at least about 110 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 130 nucleotides in length, alternatively at least about 140 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 160 nucleotides in length, alternatively at least about 170 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 190 nucleotides in length, alternatively at least about 200 nucleotides in length, alternatively at least about 250 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 350 nucleotides in length, alternatively at least about 400 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 500 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 700 nucleotides in length, alternatively at least about 800 nucleotides in length, alternatively at least about 900 nucleotides in length and alternatively at least about 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted that novel fragments of a PRO polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the PRO polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such PRO polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the PRO polypeptide fragments encoded by these nucleotide molecule fragments, preferably those PRO polypeptide fragments that comprise a binding site for an anti-PRO antibody.

In another embodiment, the invention provides an isolated PRO polypeptide encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a certain aspect, the invention concerns an isolated PRO polypeptide, comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least

about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein.

In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

Another aspect of the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

In yet another embodiment, the invention concerns agonists of a native PRO polypeptide as defined herein.

In a particular embodiment, the agonist is an anti-PRO antibody or a small molecule.

In a further embodiment, the invention concerns a method of identifying agonists to a PRO polypeptide which comprise contacting the PRO polypeptide with a candidate molecule and monitoring a biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native PRO polypeptide.

5 In a still further embodiment, the invention concerns a composition of matter comprising a PRO polypeptide, or an agonist of a PRO polypeptide as herein described, or an anti-PRO antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

10 Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist thereof as hereinbefore described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist thereof or an anti-PRO antibody.

15 In additional embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cells comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli*, yeast, or Baculovirus-infected insect cells. A process for producing any of the herein described polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

20 In other embodiments, the invention provides chimeric molecules comprising any of the herein described polypeptides fused to a heterologous polypeptide or amino acid sequence. Example of such chimeric molecules comprise any of the herein described polypeptides fused to an epitope tag sequence or a Fc region of an immunoglobulin.

In yet another embodiment, the invention provides an antibody which specifically binds to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, humanized antibody, antibody fragment or single-chain antibody.

25 In yet other embodiments, the invention provides oligonucleotide probes useful for isolating genomic and cDNA nucleotide sequences or as antisense probes, wherein those probes may be derived from any of the above or below described nucleotide sequences.

BRIEF DESCRIPTION OF THE DRAWINGS

30 Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a native sequence PRO943 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA52192-1369".

Figure 2 shows the amino acid sequence (SEQ ID NO:2) derived from the coding sequence of SEQ ID NO:1 shown in Figure 1.

Figure 3 shows a nucleotide sequence (SEQ ID NO:3) of a native sequence PRO1250 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA60775-1532".

35 Figure 4 shows the amino acid sequence (SEQ ID NO:4) derived from the coding sequence of SEQ ID NO:3 shown in Figure 3.

Figure 5 shows a nucleotide sequence (SEQ ID NO:5) of a native sequence PRO1337 cDNA, wherein

SEQ ID NO:5 is a clone designated herein as "DNA66672-1586".

Figure 6 shows the amino acid sequence (SEQ ID NO:6) derived from the coding sequence of SEQ ID NO:5 shown in Figure 5.

DETAILED DESCRIPTION OF THE INVENTION

5 The terms "PRO polypeptide", and "PRO" as used herein and when immediately followed by a numerical designation refer to various polypeptides, wherein the complete designation (*i.e.*, PRO/number) refers to specific polypeptide sequences as described herein. The terms "PRO/number polypeptide" and "PRO/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides and polypeptide variants (which are further defined herein). The PRO polypeptides described
10 herein may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods.

 A "native sequence PRO polypeptide" comprises a polypeptide having the same amino acid sequence as the corresponding PRO polypeptide derived from nature. Such native sequence PRO polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO
15 polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific PRO polypeptide (*e.g.*, an extracellular domain sequence), naturally-occurring variant forms (*e.g.*, alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In various embodiments of the invention, the native sequence PRO polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acid sequences shown in the accompanying figures. Start and
20 stop codons are shown in bold font and underlined in the figures. However, while the PRO polypeptide disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either upstream or downstream from the amino acid position 1 in the figures may be employed as the starting amino acid residue for the PRO polypeptides.

25 The PRO polypeptide "extracellular domain" or "ECD" refers to a form of the PRO polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a PRO polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the PRO polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type
30 of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a PRO polypeptide may contain from about 5 or fewer amino acids on either side of the transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated
35 by the present invention.

 The approximate location of the "signal peptides" of the various PRO polypeptides disclosed herein are shown in the present specification and/or the accompanying figures. It is noted, however, that the C-terminal

boundary of a signal peptide may vary, but most likely by no more than about 5 amino acids on either side of the signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (*e.g.*, Nielsen *et al.*, Prot. Eng., 10:1-6 (1997) and von Heinje *et al.*, Nucl. Acids Res., 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal sequence from a secreted polypeptide is not entirely uniform, resulting in more than one secreted species. These mature polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present invention.

"PRO polypeptide variant" means an active PRO polypeptide as defined above or below having at least about 80% amino acid sequence identity with a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Such PRO polypeptide variants include, for instance, PRO polypeptides wherein one or more amino acid residues are added, or deleted, at the N- or C-terminus of the full-length native amino acid sequence. Ordinarily, a PRO polypeptide variant will have at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, PRO variant polypeptides are at least about 10 amino acids in length, alternatively at least about 20 amino acids in length, alternatively at least about 30 amino acids in length, alternatively at least about 40 amino acids in length, alternatively at least about 50 amino acids in length, alternatively at least about 60 amino acids in length, alternatively at least about 70 amino acids in length, alternatively at least about 80 amino acids in length, alternatively at least about 90 amino acids in length, alternatively at least about 100 amino acids in length, alternatively at least about 150 amino acids in length, alternatively at least about 200 amino acids in length, alternatively at least about 300 amino acids in length, or more.

"Percent (%) amino acid sequence identity" with respect to the PRO polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in a PRO sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN, ALIGN-2 or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full-length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are obtained as described below by using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc., and the source code shown in Table 1 has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

For purposes herein, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations, Tables 2-3 demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "PRO".

Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described above using the ALIGN-2 sequence comparison computer program. However, % amino acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul *et al.*, Nucleic Acids Res., 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov>, or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set

to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

where X is the number of amino acid residues scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A.

In addition, % amino acid sequence identity may also be determined using the WU-BLAST-2 computer program (Altschul *et al.*, *Methods in Enzymology*, 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, *i.e.*, the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. For purposes herein, a % amino acid sequence identity value is determined by dividing (a) the number of matching identical amino acids residues between the amino acid sequence of the PRO polypeptide of interest having a sequence derived from the native PRO polypeptide and the comparison amino acid sequence of interest (*i.e.*, the sequence against which the PRO polypeptide of interest is being compared which may be a PRO variant polypeptide) as determined by WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest. For example, in the statement "a polypeptide comprising an amino acid sequence A which has or having at least 80% amino acid sequence identity to the amino acid sequence B", the amino acid sequence A is the comparison amino acid sequence of interest and the amino acid sequence B is the amino acid sequence of the PRO polypeptide of interest.

"PRO variant polynucleotide" or "PRO variant nucleic acid sequence" means a nucleic acid molecule which encodes an active PRO polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, a PRO variant polynucleotide will have at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively

at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal sequence, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

Ordinarily, PRO variant polynucleotides are at least about 30 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 210 nucleotides in length, alternatively at least about 240 nucleotides in length, alternatively at least about 270 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 900 nucleotides in length, or more.

"Percent (%) nucleic acid sequence identity" with respect to the PRO polypeptide-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in a PRO polypeptide-encoding nucleic acid sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN, ALIGN-2 or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full-length of the sequences being compared. For purposes herein, however, % nucleic acid sequence identity values are obtained as described below by using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc., and the source code shown in Table 1 has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

For purposes herein, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or

against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$100 \text{ times the fraction } W/Z$$

5 where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 4-5 demonstrate how to calculate the
10 % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "PRO-DNA".

Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described above using the ALIGN-2 sequence comparison computer program. However, % nucleic acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul
15 *et al.*, Nucleic Acids Res., 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov>, or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped
20 alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$25 \quad 100 \text{ times the fraction } W/Z$$

where W is the number of nucleotides scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity
30 of D to C.

In addition, % nucleic acid sequence identity values may also be generated using the WU-BLAST-2 computer program (Altschul *et al.*, Methods in Enzymology, 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, *i.e.*, the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and

scoring matrix = BLOSUM62. For purposes herein, a % nucleic acid sequence identity value is determined by dividing (a) the number of matching identical nucleotides between the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest having a sequence derived from the native sequence PRO polypeptide-encoding nucleic acid and the comparison nucleic acid molecule of interest (*i.e.*, the sequence
5 against which the PRO polypeptide-encoding nucleic acid molecule of interest is being compared which may be a variant PRO polynucleotide) as determined by WU-BLAST-2 by (b) the total number of nucleotides of the PRO polypeptide-encoding nucleic acid molecule of interest. For example, in the statement "an isolated nucleic acid molecule comprising a nucleic acid sequence A which has or having at least 80% nucleic acid sequence identity to the nucleic acid sequence B", the nucleic acid sequence A is the comparison nucleic acid molecule
10 of interest and the nucleic acid sequence B is the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest.

In other embodiments, PRO variant polynucleotides are nucleic acid molecules that encode an active PRO polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding the full-length PRO polypeptide shown in the accompanying figures herein.
15 PRO variant polypeptides may be those that are encoded by a PRO variant polynucleotide.

"Isolated", when used to describe the various polypeptides disclosed herein, means a polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Preferably, the isolated polypeptide is free of association with all components with which it is naturally associated. Contaminant components of its natural environment are materials that would typically interfere with diagnostic
20 or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptides includes polypeptides *in situ* within recombinant cells, since at
25 least one component of the PRO polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptides will be prepared by at least one purification step.

An "isolated" nucleic acid molecule encoding a PRO polypeptide or an "isolated" nucleic acid molecule encoding an anti-PRO antibody is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the PRO-
30 encoding nucleic acid or the natural source of the anti-PRO-encoding nucleic acid. Preferably, the isolated nucleic acid is free of association with all components with which it is naturally associated. An isolated PRO-encoding nucleic acid molecule or an isolated anti-PRO-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated nucleic acid molecules therefore are distinguished from the PRO-encoding nucleic acid molecule or from the anti-PRO-encoding nucleic acid molecule as it exists in natural
35 cells. However, an isolated nucleic acid molecule encoding a PRO polypeptide or an isolated nucleic acid molecule encoding an anti-PRO antibody includes PRO-nucleic acid molecules or anti-PRO-nucleic acid molecules contained in cells that ordinarily express PRO polypeptides or anti-PRO antibodies where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

5 Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a PRO polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally,
10 "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

15 The term "antibody" is used in the broadest sense and specifically covers, for example, single anti-PRO monoclonal antibodies (including agonist antibodies), anti-PRO antibody compositions with polypeptopic specificity, single chain anti-PRO antibodies, and fragments of anti-PRO antibodies (*see below*). The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, *i.e.*, the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

20 "Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree
25 of desired homology between the probe and hybridizable sequence, the higher the relative temperature that can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, *see, Ausubel et al., Current Protocols in Molecular Biology* (Wiley Interscience Publishers, 1995).

30 "Stringent conditions" or "high-stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example, 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium
35 citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x

SSC containing EDTA at 55°C.

"Moderately-stringent conditions" may be identified as described by Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Press, 1989), and include the use of washing solution and hybridization conditions (*e.g.*, temperature, ionic strength, and % SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37°-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a PRO polypeptide fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (*i.e.*, is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

"Active" or "activity" for the purposes herein refers to form(s) of PRO polypeptides which retain a biological and/or an immunological activity of native or naturally-occurring PRO polypeptides, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring PRO polypeptide other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO polypeptide and an "immunological" activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO polypeptide.

"Biological activity" in the context of an antibody or another agonist that can be identified by the screening assays disclosed herein (*e.g.*, an organic or inorganic small molecule, peptide, etc.) is used to refer to the ability of such molecules to invoke one or more of the effects listed herein in connection with the definition of a "therapeutically effective amount." In a specific embodiment, "biological activity" is the ability to inhibit neoplastic cell growth or proliferation. A preferred biological activity is inhibition, including slowing or complete stopping, of the growth of a target tumor (*e.g.*, cancer) cell. Another preferred biological activity is

cytotoxic activity resulting in the death of the target tumor (*e.g.*, cancer) cell. Yet another preferred biological activity is the induction of apoptosis of a target tumor (*e.g.*, cancer) cell.

The phrase "immunological activity" means immunological cross-reactivity with at least one epitope of a PRO polypeptide.

5 "Immunological cross-reactivity" as used herein means that the candidate polypeptide is capable of competitively inhibiting the qualitative biological activity of a PRO polypeptide having this activity with polyclonal antisera raised against the known active PRO polypeptide. Such antisera are prepared in conventional fashion by injecting goats or rabbits, for example, subcutaneously with the known active analogue in complete Freund's adjuvant, followed by booster intraperitoneal or subcutaneous injection in incomplete
10 Freund's. The immunological cross-reactivity preferably is "specific", which means that the binding affinity of the immunologically cross-reactive molecule (*e.g.*, antibody) identified, to the corresponding PRO polypeptide is significantly higher (preferably at least about 2-times, more preferably at least about 4-times, even more preferably at least about 6-times, most preferably at least about 8-times higher) than the binding affinity of that molecule to any other known native polypeptide.

15 "Tumor", as used herein, refers to all neoplastic cell growth and proliferation, whether malignant or benign, and all pre-cancerous and cancerous cells and tissues.

The terms "cancer" and "cancerous" refer to or describe the physiological condition in mammals that is typically characterized by unregulated cell growth. Examples of cancer include but are not limited to, carcinoma, lymphoma, blastoma, sarcoma, and leukemia. More particular examples of such cancers include
20 breast cancer, prostate cancer, colon cancer, squamous cell cancer, small-cell lung cancer, non-small cell lung cancer, ovarian cancer, cervical cancer, gastrointestinal cancer, pancreatic cancer, glioblastoma, liver cancer, bladder cancer, hepatoma, colorectal cancer, endometrial carcinoma, salivary gland carcinoma, kidney cancer, vulval cancer, thyroid cancer, hepatic carcinoma and various types of head and neck cancer.

"Treatment" is an intervention performed with the intention of preventing the development or altering the
25 pathology of a disorder. Accordingly, "treatment" refers to both therapeutic treatment and prophylactic or preventative measures. Those in need of treatment include those already with the disorder as well as those in which the disorder is to be prevented. In tumor (*e.g.*, cancer) treatment, a therapeutic agent may directly decrease the pathology of tumor cells, or render the tumor cells more susceptible to treatment by other therapeutic agents, *e.g.*, radiation and/or chemotherapy.

30 The "pathology" of cancer includes all phenomena that compromise the well-being of the patient. This includes, without limitation, abnormal or uncontrollable cell growth, metastasis, interference with the normal functioning of neighboring cells, release of cytokines or other secretory products at abnormal levels, suppression or aggravation of inflammatory or immunological response, etc.

35 An "effective amount" of a polypeptide disclosed herein or an agonist thereof, in reference to inhibition of neoplastic cell growth, is an amount capable of inhibiting, to some extent, the growth of target cells. The term includes an amount capable of invoking a growth inhibitory, cytostatic and/or cytotoxic effect and/or apoptosis of the target cells. An "effective amount" of a PRO polypeptide or an agonist thereof for purposes of inhibiting neoplastic cell growth may be determined empirically and in a routine manner.

5 A "therapeutically effective amount", in reference to the treatment of tumor, refers to an amount capable of invoking one or more of the following effects: (1) inhibition, to some extent, of tumor growth, including, slowing down and complete growth arrest; (2) reduction in the number of tumor cells; (3) reduction in tumor size; (4) inhibition (*i.e.*, reduction, slowing down or complete stopping) of tumor cell infiltration into peripheral organs; (5) inhibition (*i.e.*, reduction, slowing down or complete stopping) of metastasis; (6) enhancement of anti-tumor immune response, which may, but does not have to, result in the regression or rejection of the tumor; and/or (7) relief, to some extent, of one or more symptoms associated with the disorder. A "therapeutically effective amount" of a PRO polypeptide or an agonist thereof for purposes of treatment of tumor may be determined empirically and in a routine manner.

10 A "growth inhibitory amount" of a PRO polypeptide or an agonist thereof is an amount capable of inhibiting the growth of a cell, especially tumor, *e.g.*, cancer cell, either *in vitro* or *in vivo*. A "growth inhibitory amount" of a PRO polypeptide or an agonist thereof for purposes of inhibiting neoplastic cell growth may be determined empirically and in a routine manner.

15 A "cytotoxic amount" of a PRO polypeptide or an agonist thereof is an amount capable of causing the destruction of a cell, especially tumor, *e.g.*, cancer cell, either *in vitro* or *in vivo*. A "cytotoxic amount" of a PRO polypeptide or an agonist thereof for purposes of inhibiting neoplastic cell growth may be determined empirically and in a routine manner.

20 The term "cytotoxic agent" as used herein refers to a substance that inhibits or prevents the function of cells and/or causes destruction of cells. The term is intended to include radioactive isotopes (*e.g.*, I^{131} , I^{125} , Y^{90} and Re^{186}), chemotherapeutic agents, and toxins such as enzymatically active toxins of bacterial, fungal, plant or animal origin, or fragments thereof.

25 A "chemotherapeutic agent" is a chemical compound useful in the treatment of tumor, *e.g.*, cancer. Examples of chemotherapeutic agents include adriamycin, doxorubicin, epirubicin, 5-fluorouracil, cytosine arabinoside ("Ara-C"), cyclophosphamide, thiotepa, busulfan, cytoxan, taxoids, *e.g.*, paclitaxel (Taxol, Bristol-Myers Squibb Oncology, Princeton, NJ), and doxetaxel (Taxotere, Rhône-Poulenc Rorer, Antony, France),
30 toxotere, methotrexate, cisplatin, melphalan, vinblastine, bleomycin, etoposide, ifosfamide, mitomycin C, mitoxantrone, vincristine, vinorelbine, carboplatin, teniposide, daunomycin, carminomycin, aminopterin, dactinomycin, mitomycins, esperamicins (*see*, U.S. Patent No. 4,675,187), melphalan and other related nitrogen mustards. Also included in this definition are hormonal agents that act to regulate or inhibit hormone action on tumors such as tamoxifen and onapristone.

35 A "growth inhibitory agent" when used herein refers to a compound or composition which inhibits growth of a cell, especially tumor, *e.g.*, cancer cell, either *in vitro* or *in vivo*. Thus, the growth inhibitory agent is one which significantly reduces the percentage of the target cells in S phase. Examples of growth inhibitory agents include agents that block cell cycle progression (at a place other than S phase), such as agents that induce G1 arrest and M-phase arrest. Classical M-phase blockers include the vincas (vincristine and vinblastine), taxol, and topo II inhibitors such as doxorubicin, epirubicin, daunorubicin, etoposide, and bleomycin. Those agents that arrest G1 also spill over into S-phase arrest, for example, DNA alkylating agents such as tamoxifen, prednisone, dacarbazine, mechlorethamine, cisplatin, methotrexate, 5-fluorouracil, and ara-C. Further

information can be found in The Molecular Basis of Cancer, Mendelsohn and Israel, eds., Chapter 1, entitled "Cell cycle regulation, oncogens, and antineoplastic drugs" by Murakami *et al.*, (WB Saunders: Philadelphia, 1995), especially p. 13.

5 The term "cytokine" is a generic term for proteins released by one cell population which act on another cell as intercellular mediators. Examples of such cytokines are lymphokines, monokines, and traditional polypeptide hormones. Included among the cytokines are growth hormone such as human growth hormone, N-methionyl human growth hormone, and bovine growth hormone; parathyroid hormone; thyroxine; insulin; proinsulin; relaxin; prolaxin; glycoprotein hormones such as follicle stimulating hormone (FSH), thyroid stimulating hormone (TSH), and luteinizing hormone (LH); hepatic growth factor; fibroblast growth factor; 10 prolactin; placental lactogen; tumor necrosis factor- α and - β ; mullerian-inhibiting substance; mouse gonadotropin-associated peptide; inhibin; activin; vascular endothelial growth factor; integrin; thrombopoietin (TPO); nerve growth factors such as NGF- β ; platelet-growth factor; transforming growth factors (TGFs) such as TGF- α and TGF- β ; insulin-like growth factor-I and -II; erythropoietin (EPO); osteoinductive factors; interferons such as interferon- α , - β , and - γ ; colony stimulating factors (CSFs) such as macrophage-CSF (M- 15 CSF); granulocyte-macrophage-CSF (GM-CSF); and granulocyte-CSF (G-CSF); interleukins (ILs) such as IL-1, IL-1 α , IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12; a tumor necrosis factor such as TNF- α or TNF- β ; and other polypeptide factors including LIF and kit ligand (KL). As used herein, the term cytokine includes proteins from natural sources or from recombinant cell culture and biologically active equivalents of the native sequence cytokines.

20 The term "prodrug" as used in this application refers to a precursor or derivative form of a pharmaceutically active substance that is less cytotoxic to tumor cells compared to the parent drug and is capable of being enzymatically activated or converted into the more active parent form. See, e.g., Wilman, "Prodrugs in Cancer Chemotherapy", Biochemical Society Transactions, 14, pp. 375-382, 615th Meeting Belfast (1986) and Stella *et al.*, "Prodrugs: A Chemical Approach to Targeted Drug Delivery," Directed Drug Delivery, 25 Borchardt *et al.*, (ed.), pp. 247-267, Humana Press (1985). The prodrugs of this invention include, but are not limited to, phosphate-containing prodrugs, thiophosphate-containing prodrugs, glycosylated prodrugs or optionally substituted phenylacetamide-containing prodrugs, 5-fluorocytosine and other 5-fluorouridine prodrugs which can be derivatized into a prodrug form for use in this invention include, but are not limited to, those chemotherapeutic agents described above.

30 The term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native PRO polypeptide disclosed herein. Suitable agonist molecules specifically include agonist antibodies or antibody fragments, fragments or amino acid sequence variants of native PRO polypeptides, peptides, small organic molecules, etc. Methods for identifying agonists of a PRO polypeptide may comprise contacting a tumor cell with a candidate agonist molecule and measuring the inhibition of tumor cell growth.

35 "Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. "Intermittent" administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

"Mammal" for purposes of treatment refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is human.

Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEENTM, polyethylene glycol (PEG), and PLURONICSTM.

"Native antibodies" and "native immunoglobulins" are usually heterotetrameric glycoproteins of about 150,000 daltons, composed of two identical light (L) chains and two identical heavy (H) chains. Each light chain is linked to a heavy chain by one covalent disulfide bond, while the number of disulfide linkages varies among the heavy chains of different immunoglobulin isotypes. Each heavy and light chain also has regularly spaced intrachain disulfide bridges. Each heavy chain has at one end a variable domain (V_H) followed by a number of constant domains. Each light chain has a variable domain at one end (V_L) and a constant domain at its other end; the constant domain of the light chain is aligned with the first constant domain of the heavy chain, and the light-chain variable domain is aligned with the variable domain of the heavy chain. Particular amino acid residues are believed to form an interface between the light- and heavy-chain variable domains.

The term "variable" refers to the fact that certain portions of the variable domains differ extensively in sequence among antibodies and are used in the binding and specificity of each particular antibody for its particular antigen. However, the variability is not evenly distributed throughout the variable domains of antibodies. It is concentrated in three segments called complementarity-determining regions (CDRs) or hypervariable regions both in the light-chain and the heavy-chain variable domains. The more highly conserved portions of variable domains are called the framework regions (FR). The variable domains of native heavy and light chains each comprise four FR regions, largely adopting a β -sheet configuration, connected by three CDRs, which form loops connecting, and in some cases forming part of, the β -sheet structure. The CDRs in each chain are held together in close proximity by the FR regions and, with the CDRs from the other chain, contribute to the formation of the antigen-binding site of antibodies (*see*, Kabat *et al.*, NIH Publ. No.91-3242, Vol. I, pages 647-669 (1991)). The constant domains are not involved directly in binding an antibody to an antigen, but exhibit various effector functions, such as participation of the antibody in antibody-dependent cellular toxicity.

The term "hypervariable region" when used herein refers to the amino acid residues of an antibody which are responsible for antigen-binding. The hypervariable region comprises amino acid residues from a

"complementarity determining region" or "CDR" (*i.e.*, residues 24-34 (L1), 50-56 (L2) and 89-97 (L3) in the light chain variable domain and 31-35 (H1), 50-65 (H2) and 95-102 (H3) in the heavy chain variable domain; Kabat *et al.*, Sequences of Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institute of Health, Bethesda, MD. [1991]) and/or those residues from a "hypervariable loop" (*i.e.*, residues 26-32 (L1), 50-52 (L2) and 91-96 (L3) in the light chain variable domain and 26-32 (H1), 53-55 (H2) and 96-101 (H3) in the heavy chain variable domain; Clothia and Lesk, J. Mol. Biol., 196:901-917 [1987]). "Framework" or "FR" residues are those variable domain residues other than the hypervariable region residues as herein defined.

"Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')₂, and Fv fragments; diabodies; linear antibodies (Zapata *et al.*, Protein Eng., 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an F(ab')₂ fragment that has two antigen-combining sites and is still capable of cross-linking antigen.

"Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the V_H-V_L dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')₂ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), *e.g.*, IgG1, IgG2, IgG3, IgG4, IgA, and IgA2.

The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, *i.e.*, the individual antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies

are highly specific, being directed against a single antigenic site. Furthermore, in contrast to conventional (polyclonal) antibody preparations which typically include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. In addition to their specificity, the monoclonal antibodies are advantageous in that they are synthesized by the hybridoma culture, uncontaminated by other immunoglobulins. The modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies to be used in accordance with the present invention may be made by the hybridoma method first described by Kohler *et al.*, Nature, 256:495 [1975], or may be made by recombinant DNA methods (see, e.g., U.S. Patent No. 4,816,567). The "monoclonal antibodies" may also be isolated from phage antibody libraries using the techniques described in Clackson *et al.*, Nature, 352:624-628 [1991] and Marks *et al.*, J. Mol. Biol., 222:581-597 (1991), for example.

The monoclonal antibodies herein specifically include "chimeric" antibodies (immunoglobulins) in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity (U.S. Patent No. 4,816,567; Morrison *et al.*, Proc. Natl. Acad. Sci. USA, 81:6851-6855 [1984]).

"Humanized" forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a CDR of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity, and capacity. In some instances, Fv FR residues of the human immunoglobulin are replaced by corresponding non-human residues. Furthermore, humanized antibodies may comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. These modifications are made to further refine and maximize antibody performance. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. For further details, see, Jones *et al.*, Nature, 321:522-525 (1986); Reichmann *et al.*, Nature, 332:323-329 [1988]; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992). The humanized antibody includes a PRIMATIZED™ antibody wherein the antigen-binding region of the antibody is derived from an antibody produced by immunizing macaque monkeys with the antigen of interest.

"Single-chain Fv" or "sFv" antibody fragments comprise the V_H and V_L domains of antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a

polypeptide linker between the V_H and V_L domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, *see*, Pluckthun in The Pharmacology of Monoclonal Antibodies, Vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

5 The term "diabodies" refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) in the same polypeptide chain ($V_H - V_L$). By using a linker that is too short to allow pairing between the two domains on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger *et al.*, Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

10 An "isolated" antibody is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than
15 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody *in situ* within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

20 An antibody that "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide is one that binds to that particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody so as to generate a "labeled" antibody. The label may be detectable by itself
25 (*e.g.*, radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable. The label may also be a non-detectable entity such as a toxin.

By "solid phase" is meant a non-aqueous matrix to which the antibody of the present invention can adhere. Examples of solid phases encompassed herein include those formed partially or entirely of glass (*e.g.*, controlled
30 pore glass), polysaccharides (*e.g.*, agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (*e.g.*, an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which
35 is useful for delivery of a drug (such as a PRO polypeptide or antibody thereto) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

A "small molecule" is defined herein to have a molecular weight below about 500 Daltons.

As shown below, Table 1 provides the complete source code for the ALIGN-2 sequence comparison computer program. This source code may be routinely compiled for use on a UNIX operating system to provide the ALIGN-2 sequence comparison computer program.

5 In addition, Tables 2-5 show hypothetical exemplifications for using the above described method to determine % amino acid sequence identity (Tables 2-3) and % nucleic acid sequence identity (Tables 4-5) using the ALIGN-2 sequence comparison computer program, wherein "PRO" represents the amino acid sequence of a hypothetical PRO polypeptide of interest, "Comparison Protein" represents the amino acid sequence of a polypeptide against which the "PRO" polypeptide of interest is being compared, "PRO-DNA" represents a
10 hypothetical PRO-encoding nucleic acid sequence of interest, "Comparison DNA" represents the nucleotide sequence of a nucleic acid molecule against which the "PRO-DNA" nucleic acid molecule of interest is being compared, "X", "Y", and "Z" each represent different hypothetical amino acid residues and "N", "L" and "V" each represent different hypothetical nucleotides.

Table 1

```

/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 / * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define      _M      -8      /* value of a match with a stop */

int  _day[26][26] = {
/*      A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */      { 2, 0, -2, 0, 0, -4, 1, -1, -1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */      { 0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, 0, -2, -5, 0, -3, 1},
/* C */      {-2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, -4, _M, -3, -5, -4, 0, -2, 0, -2, -8, 0, 0, -5},
/* D */      { 0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */      { 0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */      {-4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, -3, 0, -1, 0, 0, 7, -5},
/* G */      { 1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, -1, -7, 0, -5, 0},
/* H */      {-1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, -1, 0, -2, -3, 0, 0, 2},
/* I */      {-1, -2, -2, -2, -2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */      { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */      {-1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
/* L */      {-2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */      {-1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, -1},
/* N */      { 0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */      {_M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M},
/* P */      { 1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, -1, -6, 0, -5, 0},
/* Q */      { 0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, -1, 0, -2, -5, 0, -4, 3},
/* R */      {-2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
/* S */      { 1, 0, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
/* T */      { 1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 3, 0, 0, -5, 0, -3, 0},
/* U */      { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */      { 0, -2, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
/* W */      {-6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
/* X */      { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */      {-3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -5, -4, -4, -3, -3, 0, -2, 0, 0, 10, -4},
/* Z */      { 0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, 0, -2, -6, 0, -4, 4}
};

```

Table 1 (cont')

```

/*
*/
#include <stdio.h>
#include <ctype.h>

#define MAXJMP 16 /* max jumps in a diag */
#define MAXGAP 24 /* don't continue to penalize gaps larger than this */
#define JMPS 1024 /* max jmps in an path */
#define MX 4 /* save if there's at least MX-1 bases since last jmp */

#define DMAT 3 /* value of matching bases */
#define DMIS 0 /* penalty for mismatched bases */
#define DINS0 8 /* penalty for a gap */
#define DINS1 1 /* penalty per base */
#define PINS0 8 /* penalty for a gap */
#define PINS1 4 /* penalty per residue */

struct jmp {
    short n[MAXJMP]; /* size of jmp (neg for dely) */
    unsigned short x[MAXJMP]; /* base no. of jmp in seq x */
}; /* limits seq to 2^16 -1 */

struct diag {
    int score; /* score at last jmp */
    long offset; /* offset of prev block */
    short ijmp; /* current jmp index */
    struct jmp jp; /* list of jmps */
};

struct path {
    int spc; /* number of leading spaces */
    short n[JMPS]; /* size of jmp (gap) */
    int x[JMPS]; /* loc of jmp (last elem before gap) */
};

char *ofile; /* output file name */
char *namex[2]; /* seq names: getseqs() */
char *prog; /* prog name for err msgs */
char *seqx[2]; /* seqs: getseqs() */
int dmax; /* best diag: nw() */
int dmax0; /* final diag */
int dna; /* set if dna: main() */
int endgaps; /* set if penalizing end gaps */
int gapx, gapy; /* total gaps in seqs */
int len0, len1; /* seq lens */
int ngapx, ngapy; /* total size of gaps */
int smax; /* max score: nw() */
int *xbm; /* bitmap for matching */
long offset; /* current offset in jmp file */
struct diag *dx; /* holds diagonals */
struct path pp[2]; /* holds path for seqs */

char *calloc(), *malloc(), *index(), *strcpy();
char *getseq(), *g_calloc();

```

Table 1 (cont')

```

/* Needleman-Wunsch alignment program
*
* usage: progs file1 file2
* where file1 and file2 are two dna or two protein sequences.
* The sequences can be in upper- or lower-case and may contain ambiguity
* Any lines beginning with ';', '>' or '<' are ignored
* Max file length is 65535 (limited by unsigned short x in the jmp struct)
* A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
* Output is in the file "align.out"
*
* The program may create a tmp file in /tmp to hold info about traceback.
* Original version developed under BSD 4.3 on a vax 8650
*/
#include "nw.h"
#include "day.h"

static _dbval[26] = {
    1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static _pbval[26] = {
    1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
    128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
    1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
    1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)
int    ac;
char   *av[ ];
{
    prog = av[0];
    if (ac != 3) {
        fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
    }
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dna)? _dbval : _pbval;

    endgaps = 0;                /* 1 to penalize endgaps */
    ofile = "align.out";        /* output file */

    nw();                       /* fill in the matrix, get the possible jmps */
    readjmps();                 /* get the actual jmps */
    print();                    /* print stats, alignment */

    cleanup(0);                 /* unlink any tmp files */
}

```

main

Table 1 (cont')

```

/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw()
{
    char      *px, *py;          /* seqs and ptrs */
    int       *ndely, *dely;     /* keep track of dely */
    int       ndelx, delx;       /* keep track of delx */
    int       *tmp;              /* for swapping row0, row1 */
    int       mis;               /* score for each type */
    int       ins0, ins1;        /* insertion penalties */
    register  id;                /* diagonal index */
    register  ij;                /* jmp index */
    register  *col0, *col1;      /* score for curr, last row */
    register  xx, yy;            /* index into seqs */

    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
    col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;

    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0;          /* Waterman Bull Math Biol 84 */
    }
    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;

    /* fill in match matrix
    */
    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
        */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }
}

```

nw

Table 1 (cont')

...nw

```

for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
    mis = col0[yy-1];
    if (dna)
        mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
    else
        mis += _day[*px-'A'][*py-'A'];

    /* update penalty for del in x seq;
    * favor new del over ongong del
    * ignore MAXGAP if weighting endgaps
    */
    if (endgaps || ndely[yy] < MAXGAP) {
        if (col0[yy] - ins0 >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0 + ins1);
            ndely[yy] = 1;
        } else {
            dely[yy] -= ins1;
            ndely[yy]++;
        }
    } else {
        if (col0[yy] - (ins0 + ins1) >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0 + ins1);
            ndely[yy] = 1;
        } else
            ndely[yy]++;
    }

    /* update penalty for del in y seq;
    * favor new del over ongong del
    */
    if (endgaps || ndelx < MAXGAP) {
        if (col1[yy-1] - ins0 >= delx) {
            delx = col1[yy-1] - (ins0 + ins1);
            ndelx = 1;
        } else {
            delx -= ins1;
            ndelx++;
        }
    } else {
        if (col1[yy-1] - (ins0 + ins1) >= delx) {
            delx = col1[yy-1] - (ins0 + ins1);
            ndelx = 1;
        } else
            ndelx++;
    }

    /* pick the maximum score; we're favoring
    * mis over any del and delx over dely
    */

```

Table 1 (cont')

...nw

```

id = xx - yy + len1 - 1;
if (mis >= delx && mis >= dely[yy])
    col1[yy] = mis;
else if (delx >= dely[yy]) {
    col1[yy] = delx;
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writejumps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
    }
    dx[id].jp.n[ij] = ndelx;
    dx[id].jp.x[ij] = xx;
    dx[id].score = delx;
}
else {
    col1[yy] = dely[yy];
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writejumps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
    }
    dx[id].jp.n[ij] = -ndely[yy];
    dx[id].jp.x[ij] = xx;
    dx[id].score = dely[yy];
}
if (xx == len0 && yy < len1) {
    /* last col
    */
    if (endgaps)
        col1[yy] -= ins0+ins1*(len1-yy);
    if (col1[yy] > smax) {
        smax = col1[yy];
        dmax = id;
    }
}
}
if (endgaps && xx < len0)
    col1[yy-1] -= ins0+ins1*(len0-xx);
if (col1[yy-1] > smax) {
    smax = col1[yy-1];
    dmax = id;
}
tmp = col0; col0 = col1; col1 = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
(void) free((char *)col0);
(void) free((char *)col1);
}

```

Table 1 (cont')

```

/*
 *
 * print() -- only routine visible outside this module
 *
 * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[ ]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() -- put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC 3
#define P__LINE 256 /* maximum output line */
#define P__SPC 3 /* space between name or num and seq */

extern _day[26][26];
int olen; /* set output line length */
FILE *fx; /* output file */

print()
{
    int lx, ly, firstgap, lastgap; /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "< first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "< second sequence: %s (length = %d)\n", namex[1], len1);
    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

```

print

Table 1 (cont')

```

/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)                                getmat
{
    int      lx, ly;      /* "core" (minus endgaps) */
    int      firstgap, lastgap;  /* leading trailing overlap */

    int      nm, i0, i1, siz0, siz1;
    char      outx[32];
    double    pct;
    register  n0, n1;
    register char *p0, *p1;

    /* get total matches, score
    */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
    n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

    nm = 0;
    while ( *p0 && *p1 ) {
        if (siz0) {
            p1++;
            n1++;
            siz0--;
        }
        else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
        else {
            if (xbm[*p0-'A']&xbm[*p1-'A'])
                nm++;
            if (n0++ == pp[0].x[i0])
                siz0 = pp[0].n[i0++];
            if (n1++ == pp[1].x[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
        }
    }

    /* pct homology:
    * if penalizing endgaps, base is the shorter seq
    * else, knock off overhangs and take shorter core
    */
    if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
        lx = (lx < ly)? lx : ly;
    pct = 100.*(double)nm/(double)lx;
    fprintf(fx, "\n");
    fprintf(fx, "< %d match %s in an overlap of %d: %.2f percent similarity\n",
        nm, (nm == 1)? "" : "es", lx, pct);

```

Table 1 (cont')

```

fprintf(fx, "< gaps in first sequence: %d", gapx);
if (gapx) {
    (void) sprintf(outx, "(%d %s%s)",
        ngapx, (dna)? "base":"residue", (ngapx == 1)? ":" : "s");
    fprintf(fx, "%s", outx);

    fprintf(fx, ", gaps in second sequence: %d", gapy);
    if (gapy) {
        (void) sprintf(outx, "(%d %s%s)",
            ngapy, (dna)? "base":"residue", (ngapy == 1)? ":" : "s");
        fprintf(fx, "%s", outx);
    }
    if (dna)
        fprintf(fx,
            "\n< score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
            smax, DMAT, DMIS, DINS0, DINS1);
    else
        fprintf(fx,
            "\n< score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
            smax, PINS0, PINS1);
    if (endgaps)
        fprintf(fx,
            "< endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
            firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
            lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
    else
        fprintf(fx, "< endgaps not penalized\n");
}

static          nm;          /* matches in core -- for checking */
static          lmax;        /* lengths of stripped file names */
static          ij[2];       /* jmp index for a path */
static          nc[2];       /* number at start of current line */
static          ni[2];       /* current elem number -- for gapping */
static          siz[2];
static char      *ps[2];      /* ptr to current element */
static char      *po[2];      /* ptr to next output char slot */
static char      out[2][P_LINE]; /* output line */
static char      star[P_LINE]; /* set by stars() */

/*
 * print alignment of described in struct path pp[ ]
 */
static
pr_align()
{
    int          nn;          /* char count */
    int          more;
    register     i;

    for (i = 0, lmax = 0; i < 2; i++) {
        nn = stripname(name[i]);
        if (nn > lmax)
            lmax = nn;

        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
        ps[i] = seqx[i];
        po[i] = out[i];
    }
}

```

...getmat

pr_align

Table 1 (cont')

```

for (nn = nm = 0, more = 1; more;) {
    for (i = more = 0; i < 2; i++) {
        /*
         * do we have more of this sequence?
         */
        if (!*ps[i])
            continue;

        more++;

        if (pp[i].spc) { /* leading space */
            *po[i]++ = ' ';
            pp[i].spc--;
        }
        else if (siz[i]) { /* in a gap */
            *po[i]++ = '-';
            siz[i]--;
        }
        else {
            /* we're putting a seq element
             */
            *po[i] = *ps[i];
            if (islower(*ps[i]))
                *ps[i] = toupper(*ps[i]);
            po[i]++;
            ps[i]++;

            /*
             * are we at next gap for this seq?
             */
            if (ni[i] == pp[i].x[ij[i]]) {
                /*
                 * we need to merge all gaps
                 * at this location
                 */
                siz[i] = pp[i].n[ij[i] + +];
                while (ni[i] == pp[i].x[ij[i]])
                    siz[i] += pp[i].n[ij[i] + +];
            }
            ni[i]++;
        }
    }
    if (++nn == olen || !more && nn) {
        dumpblock();
        for (i = 0; i < 2; i++)
            po[i] = out[i];
        nn = 0;
    }
}

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
static
dumpblock()
{
    register i;

    for (i = 0; i < 2; i++)
        *po[i]-- = '\0';
}

```

...pr_align

dumpblock

Table 1 (cont')

...dumpblock

```

(void) putc('\n', fx);
for (i = 0; i < 2; i++) {
    if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
        if (i == 0)
            nums(i);
        if (i == 0 && *out[1])
            stars();
        putline(i);
        if (i == 0 && *out[1])
            fprintf(fx, star);
        if (i == 1)
            nums(i);
    }
}
}

```

```

/*
 * put out a number line: dumpblock()
 */
static

```

```

nums(ix)
{
    int    ix;        /* index in out[ ] holding seq line */
    char    nline[P_LINE];
    register i, j;
    register char *pn, *px, *py;

    for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
        *pn = ' ';
    for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
        else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -i : i;
                for (px = pn; j /= 10, px--)
                    *px = j%10 + '0';
                if (i < 0)
                    *px = '-';
            }
            else
                *pn = ' ';
            i++;
        }
    }
    *pn = '\0';
    nc[ix] = i;
    for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}

```

nums

```

/*
 * put out a line (name, [num], seq, [num]): dumpblock()
 */
static
putline(ix)
    int    ix;
    {

```

putline

Table 1 (cont')

...putline

```

int          i;
register char *px;

for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
    (void) putc(*px, fx);
for (; i < lmax+P_SPC; i++)
    (void) putc(' ', fx);

/* these count from 1:
 * ni[ ] is current element (from 1)
 * nc[ ] is number at start of current line
 */
for (px = out[ix]; *px; px++)
    (void) putc(*px&0x7F, fx);
(void) putc('\n', fx);
}

/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
stars()
{
    int          i;
    register char *p0, *p1, cx, *px;

    if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
        !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
        return;
    px = star;
    for (i = lmax+P_SPC; i; i--)
        *px++ = ' ';

    for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
        if (isalpha(*p0) && isalpha(*p1)) {
            if (xbm[*p0-'A']&xbm[*p1-'A']) {
                cx = '*';
                nm++;
            }
            else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
                cx = '.';
            else
                cx = ' ';
        }
        else
            cx = ' ';
        *px++ = cx;
    }
    *px++ = '\n';
    *px = '\0';
}

```

stars

Table 1 (cont')

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
stripname(pn)
char *pn; /* file name (may be path) */
{
    register char *px, *py;

    py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
            py = px + 1;
    if (py)
        (void) strcpy(pn, py);
    return(strlen(pn));
}
```

stripname

Table 1 (cont')

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
 * readjumps() -- get the good jumps, from tmp file if necessary
 * writejumps() -- write a filled array of jumps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

char *jname = "/tmp/homgXXXXXX";          /* tmp file for jumps */
FILE *fj;

int  cleanup();                          /* cleanup tmp file */
long lseek();

/*
 * remove any tmp file if we blow
 */
cleanup(i)                                cleanup
{
    int i;

    if (fj)
        (void) unlink(jname);
    exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char *
getseq(file, len)                          getseq
{
    char *file;      /* file name */
    int *len;        /* seq len */

    char line[1024], *pseq;
    register char *px, *py;
    int natgc, tlen;
    FILE *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';

```

Table 1 (cont')

...getseq

```

py = pseq + 4;
*len = tlen;
rewind(fp);

while (fgets(line, 1024, fp)) {
    if (*line == ';' || *line == '<' || *line == '>')
        continue;
    for (px = line; *px != '\n'; px++) {
        if (isupper(*px))
            *py++ = *px;
        else if (islower(*px))
            *py++ = toupper(*px);
        if (index("ATGCU", *(py-1)))
            natgc++;
    }
    *py++ = '\0';
    *py = '\0';
    (void) fclose(fp);
    dna = natgc > (tlen/3);
    return(pseq+4);
}

char *
g_alloc(msg, nx, sz)
char *msg;          /* program, calling routine */
int nx, sz;          /* number and size of elements */
{
    char *px, *calloc();

    if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
        if (*msg) {
            fprintf(stderr, "%s: g_alloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
            exit(1);
        }
    }
    return(px);
}

/*
 * get final jmps from dx[ ] or tmp file, set pp[ ], reset dmax: main()
 */
readjmps()
{
    int fd = -1;
    int siz, i0, i1;
    register i, j, xx;

    if (fj) {
        (void) fclose(fj);
        if ((fd = open(jname, O_RDONLY, 0)) < 0) {
            fprintf(stderr, "%s: can't open() %s\n", prog, jname);
            cleanup(1);
        }
    }

    for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; i++) {
        while (1) {
            for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                ;

```

g_alloc

readjmps

Table 1 (cont')

...readjumps

```

    if (j < 0 && dx[dmax].offset && fj) {
        (void) lseek(fd, dx[dmax].offset, 0);
        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
        dx[dmax].ijmp = MAXJMP-1;
    }
    else
        break;
}
if (i >= JMPS) {
    fprintf(stderr, "%s: too many gaps in alignment\n", prog);
    cleanup(1);
}
if (j >= 0) {
    siz = dx[dmax].jp.n[j];
    xx = dx[dmax].jp.x[j];
    dmax += siz;
    if (siz < 0) { /* gap in second seq */
        pp[1].n[i1] = -siz;
        xx += siz;
        /* id = xx - yy + len1 - 1
        */
        pp[1].x[i1] = xx - dmax + len1 - 1;
        gapy++;
        ngapy -= siz;
        /* ignore MAXGAP when doing endgaps */
        siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
        i1++;
    }
    else if (siz > 0) { /* gap in first seq */
        pp[0].n[i0] = siz;
        pp[0].x[i0] = xx;
        gapx++;
        ngapx += siz;
        /* ignore MAXGAP when doing endgaps */
        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
        i0++;
    }
}
else
    break;
}

/* reverse the order of jumps
*/
for (j = 0, i0--; j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--; j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
    (void) close(fd);
if (fj) {
    (void) unlink(jname);
    fj = 0;
    offset = 0;
}
}

```

Table 1 (cont')

```

/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
writejumps(ix)
{
    int ix;
    char *mktemp();

    if (!fj) {
        if (mktemp(jname) < 0) {
            fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
            cleanup(1);
        }
        if ((fj = fopen(jname, "w")) == 0) {
            fprintf(stderr, "%s: can't write %s\n", prog, jname);
            exit(1);
        }
    }
    (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
    (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}

```

writejumps

Table 2

PRO	XXXXXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXXXXYYYYYYY	(Length = 12 amino acids)

% amino acid sequence identity =

- 5 (the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

5 divided by 15 = 33.3%

Table 3

PRO	XXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXXYYYYYYZZYZ	(Length = 15 amino acids)

% amino acid sequence identity = .

- 5 (the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

5 divided by 10 = 50%

Table 4

PRO-DNA	NNNNNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNLLLLLLLLLL	(Length = 16 nucleotides)

% nucleic acid sequence identity =

- 5 (the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

6 divided by 14 = 42.9%

Table 5

PRO-DNA	NNNNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLTVV	(Length = 9 nucleotides)

% nucleic acid sequence identity =

- 5 (the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

4 divided by 12 = 33.3%

II. Compositions and Methods of the Invention

A. Full-length PRO Polypeptides

5 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO polypeptides. In particular, cDNAs encoding the PRO polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

As disclosed in the Examples below, cDNA clones encoding PRO polypeptides have been deposited with the ATCC. The actual nucleotide sequences of the clones can readily be determined by the skilled artisan by sequencing of the deposited clones using routine methods in the art. The predicted amino acid sequences can be determined from the nucleotide sequences using routine skill. For the PRO polypeptides and encoding nucleic acids described herein, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

B. PRO Variants

15 In addition to the full-length native sequence PRO polypeptides described herein, it is contemplated that PRO variants can be prepared. PRO variants can be prepared by introducing appropriate nucleotide changes into the PRO DNA, and/or by synthesis of the desired PRO polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO polypeptide, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

20 Variations in the native full-length sequence PRO polypeptide or in various domains of the PRO polypeptide described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO polypeptide that results in a change in the amino acid sequence of the PRO polypeptide as compared with the native sequence PRO polypeptide. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO polypeptide. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO polypeptide with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, *i.e.*, conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

30 PRO polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the PRO polypeptide.

35 PRO fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating PRO fragments by

enzymatic digestion, *e.g.*, by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, PRO polypeptide fragments share at least one biological and/or immunological activity with the native PRO polypeptide shown in the accompanying figures.

In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 6

	<u>Original Residue</u>	<u>Exemplary Substitutions</u>	<u>Preferred Substitutions</u>
15	Ala (A)	val; leu; ile	val
	Arg (R)	lys; gln; asn	lys
	Asn (N)	gln; his; lys; arg	gln
	Asp (D)	glu	glu
	Cys (C)	ser	ser
20	Gln (Q)	asn	asn
	Glu (E)	asp	asp
	Gly (G)	pro; ala	ala
	His (H)	asn; gln; lys; arg	arg
	Ile (I)	leu; val; met; ala; phe;	
25		norleucine	leu
	Leu (L)	norleucine; ile; val;	
		met; ala; phe	ile
	Lys (K)	arg; gln; asn	arg
	Met (M)	leu; phe; ile	leu
30	Phe (F)	leu; val; ile; ala; tyr	leu
	Pro (P)	ala	ala
	Ser (S)	thr	thr
	Thr (T)	ser	ser
	Trp (W)	tyr; phe	tyr
35	Tyr (Y)	trp; phe; thr; ser	phe
	Val (V)	ile; leu; met; phe;	
		ala; norleucine	leu

Substantial modifications in function or immunological identity of the PRO polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

(1) hydrophobic: norleucine, met, ala, val, leu, ile;

- (2) neutral hydrophilic: cys, ser, thr;
- (3) acidic: asp, glu;
- (4) basic: asn, gln, his, lys, arg;
- (5) residues that influence chain orientation: gly, pro; and
- 5 (6) aromatic: trp, tyr, phe.

Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

10 The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter *et al.*, Nucl. Acids Res., 13:4331 (1986); Zoller *et al.*, Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells *et al.*, Gene, 34:315 (1985)], restriction selection mutagenesis [Wells *et al.*, Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO variant DNA.

15 Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244: 1081-1085 (1989)]. Alanine is also typically
20 preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

C. Modifications of PRO Polypeptides

25 Covalent modifications of PRO polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the PRO polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO polypeptides to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO
30 antibodies, and vice-versa. Commonly used crosslinking agents include, *e.g.*, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

35 Other modifications include deamidation of glutamyl and asparaginy residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp.

79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO polypeptides (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence PRO polypeptide. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

Addition of glycosylation sites to the PRO polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO polypeptide (for O-linked glycosylation sites). The PRO polypeptide amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, *e.g.*, in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, *et al.*, Arch. Biochem. Biophys., 259:52 (1987) and by Edge *et al.*, Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura *et al.*, Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO polypeptides comprises linking the PRO polypeptide to one of a variety of nonproteinaceous polymers, *e.g.*, polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

The PRO polypeptide of the present invention may also be modified in a way to form a chimeric molecule comprising a PRO polypeptide fused to another, heterologous polypeptide or amino acid sequence.

In one embodiment, such a chimeric molecule comprises a fusion of the PRO polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO polypeptide. The presence of such epitope-tagged forms of the PRO polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-His) or poly-

histidine-glycine (poly-His-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field *et al.*, Mol. Cell. Biol., **8**:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan *et al.*, Molecular and Cellular Biology, **5**:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky *et al.*, Protein Engineering, **3**(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp *et al.*, BioTechnology, **6**:1204-1210 (1988)]; the KT3 epitope peptide [Martin *et al.*, Science, **255**:192-194 (1992)]; an α -tubulin epitope peptide [Skinner *et al.*, J. Biol. Chem., **266**:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth *et al.*, Proc. Natl. Acad. Sci. USA, **87**:6393-6397 (1990)].

In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a PRO polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG1 molecule. For the production of immunoglobulin fusions *see* also, U.S. Patent No. 5,428,130 issued June 27, 1995.

D. Preparation of PRO Polypeptides

The description below relates primarily to production of PRO polypeptides by culturing cells transformed or transfected with a vector containing PRO polypeptide nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO polypeptides. For instance, the PRO polypeptide sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [*see, e.g.*, Stewart *et al.*, Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., **85**:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the PRO polypeptide may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO polypeptide.

1. Isolation of DNA Encoding PRO Polypeptides

DNA encoding PRO polypeptides may be obtained from a cDNA library prepared from tissue believed to possess the PRO mRNA and to express it at a detectable level. Accordingly, human PRO DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO-encoding gene may also be obtained from a genomic library or by known synthetic procedures (*e.g.*, automated nucleic acid synthesis).

Libraries can be screened with probes (such as antibodies to the PRO polypeptide or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as

described in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding the PRO polypeptide is to use PCR methodology [Sambrook *et al.*, *supra*; Dieffenbach *et al.*, PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

5 The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ³²P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook *et al.*, *supra*.

10 Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

15 Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook *et al.*, *supra*, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

20 Host cells are transfected or transformed with expression or cloning vectors described herein for PRO polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook *et al.*, *supra*.

30 Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example, CaCl₂, CaPO₄, liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employed calcium chloride, as described in Sambrook *et al.*, *supra*, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw *et al.*, Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen *et al.*, J. Bact., 130:946 (1977) and Hsiao *et al.*, Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection,

electroporation, bacterial protoplast fusion with intact cells, or polycations, *e.g.*, polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, *see*, Keown *et al.*, Methods in Enzymology, 185:527-537 (1990) and Mansour *et al.*, Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, *e.g.*, *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, *e.g.*, *Salmonella typhimurium*, *Serratia*, *e.g.*, *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (*e.g.*, *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA*; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan^r*; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan^r*; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, *e.g.*, PCR or other nucleic acid polymerase reactions, are suitable.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer *et al.*, Bio/Technology, 9:968-975 (1991)) such as, *e.g.*, *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt *et al.*, J. Bacteriol., 737 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickerhamii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilum* (ATCC 36,906; Van den Berg *et al.*, Bio/Technology, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus*; *yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna *et al.*, J. Basic Microbiol., 28:265-278 [1988]); *Candida*; *Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case *et al.*, Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, *e.g.*, *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance *et al.*, Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn *et al.*, Gene, 26:205-221 [1983]; Yelton *et al.*, Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 [1985]). Methylophilic yeasts are suitable herein and include, but are not

limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylotrophs, 269 (1982).

Suitable host cells for the expression of glycosylated PRO polypeptides are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham *et al.*, J. Gen. Virol., 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (*e.g.*, cDNA or genomic DNA) encoding PRO polypeptides may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The PRO polypeptide may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, *e.g.*, the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 μ plasmid

origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, *e.g.*, the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub *et al.*, Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb *et al.*, Nature, 282:39 (1979); Kingsman *et al.*, Gene, 7:141 (1979); Tschemper *et al.*, Gene, 10:157 (1980)]. The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the PRO-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems [Chang *et al.*, Nature, 275:615 (1978); Goeddel *et al.*, Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (*trp*) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the *tac* promoter [deBoer *et al.*, Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the PRO polypeptide.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman *et al.*, J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess *et al.*, J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, *e.g.*, the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with

the host cell systems.

Transcription of a DNA encoding the PRO polypeptide by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO polypeptide coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding the PRO polypeptide.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO polypeptides in recombinant vertebrate cell culture are described in Gething *et al.*, Nature, 293:620-625 (1981); Mantei *et al.*, Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO DNA and encoding a specific antibody epitope.

5. Purification of PRO Polypeptides

Forms of PRO polypeptides may be recovered from culture medium or from host cell lysates. If

membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g., Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO polypeptides can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

5 It may be desired to purify PRO polypeptides from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal
10 chelating columns to bind epitope-tagged forms of the PRO polypeptide. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO polypeptide produced.

15 E. Antibodies

Some drug candidates for use in the compositions and methods of the present invention are antibodies and antibody fragments which mimic the biological activity of a PRO polypeptide.

1. Polyclonal Antibodies

20 Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole
25 limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

2. Monoclonal Antibodies

30 The antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

35 The immunizing agent will typically include the PRO polypeptide or a fusion protein thereof. Generally,

either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur *et al.*, Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the PRO polypeptide. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, *supra*]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (*e.g.*, by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells,

or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison *et al.*, *supra*] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Human and Humanized Antibodies

The antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (*e.g.*, murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones *et al.*, *Nature*, 321:522-525 (1986); Riechmann *et al.*, *Nature*, 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones *et al.*, *Nature*, 321:522-525 (1986); Riechmann *et al.*, *Nature*, 332:323-327 (1988); Verhoeyen *et al.*,

Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991); Marks *et al.*, *J. Mol. Biol.*, 222:581 (1991)]. The techniques of Cole *et al.*, and Boerner *et al.*, are also available for the preparation of human monoclonal antibodies (Cole *et al.*, *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, p. 77 (1985) and Boerner *et al.*, *J. Immunol.*, 147(1):86-95 (1991)]. Similarly, human antibodies can be made by the introducing of human immunoglobulin loci into transgenic animals, *e.g.*, mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks *et al.*, *Bio/Technology*, 10: 779-783 (1992); Lonberg *et al.*, *Nature*, 368: 856-859 (1994); Morrison, *Nature*, 368: 812-13 (1994); Fishwild *et al.*, *Nature Biotechnology*, 14:845-51 (1996); Neuberger, *Nature Biotechnology*, 14: 826 (1996); Lonberg and Huszar, *Intern. Rev. Immunol.*, 13 :65-93 (1995).

4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the PRO polypeptide, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, *Nature*, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, *EMBO J.*, 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the

immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies *see*, for example, Suresh *et al.*, Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (*e.g.*, tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (*e.g.*, alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (*e.g.*, F(ab')₂ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan *et al.*, Science, 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Fab' fragments may be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby *et al.*, J. Exp. Med., 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny *et al.*, J. Immunol., 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger *et al.*, Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another

strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. *See, Gruber et al., J. Immunol., 152:5368 (1994).*

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. *Tutt et al., J. Immunol., 147:60 (1991).*

5 Exemplary bispecific antibodies may bind to two different epitopes on a given PRO polypeptide herein. Alternatively, an anti-PRO polypeptide arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (*e.g.*, CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular PRO polypeptide. Bispecific antibodies may also be used to localize
10 cytotoxic agents to cells which express a particular PRO polypeptide. These antibodies possess a PRO-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the PRO polypeptide and further binds tissue factor (TF).

5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies
15 are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond.
20 Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

6. Effector Function Engineering

It may be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, *e.g.*, the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) may be
25 introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). *See, Caron et al., J. Exp. Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992).* Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in
30 Wolff *et al.*, Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. *See, Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).*

7. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent
35 such as a chemotherapeutic agent, toxin (*e.g.*, an enzymatically active toxin of bacterial, fungal, plant, or animal

origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolacca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi , ^{131}I , ^{131}In , ^{90}Y , and ^{186}Re .

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, **238**: 1098(1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. *See*, WO94/11026.

In another embodiment, the antibody may be conjugated to a "receptor" (such as streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (*e.g.*, avidin) that is conjugated to a cytotoxic agent (*e.g.*, a radionucleotide).

8. Immunoliposomes

The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein *et al.*, Proc. Natl. Acad. Sci. USA, **82**: 3688 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA, **77**: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556.

Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin *et al.*, J. Biol. Chem., **257**: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. *See*, Gabizon *et al.*, J. National Cancer Inst., **81**(19): 1484 (1989).

F. Identification of Proteins Capable of Inhibiting Neoplastic Cell Growth or Proliferation

The proteins disclosed in the present application have been assayed in a panel of 60 tumor cell lines

currently used in the investigational, disease-oriented, *in vitro* drug-discovery screen of the National Cancer Institute (NCI). The purpose of this screen is to identify molecules that have cytotoxic and/or cytostatic activity against different types of tumors. NCI screens more than 10,000 new molecules per year (Monks *et al.*, J. Natl. Cancer Inst., 83:757-766 (1991); Boyd, Cancer: Princ. Pract. Oncol. Update, 3(10):1-12 ([1989])). The tumor cell lines employed in this study have been described in Monks *et al.*, *supra*. The cell lines the growth of which has been significantly inhibited by the proteins of the present application are specified in the Examples.

The results have shown that the proteins tested show cytostatic and, in some instances and concentrations, cytotoxic activities in a variety of cancer cell lines, and therefore are useful candidates for tumor therapy.

Other cell-based assays and animal models for tumors (*e.g.*, cancers) can also be used to verify the findings of the NCI cancer screen, and to further understand the relationship between the protein identified herein and the development and pathogenesis of neoplastic cell growth. For example, primary cultures derived from tumors in transgenic animals (as described below) can be used in the cell-based assays herein, although stable cell lines are preferred. Techniques to derive continuous cell lines from transgenic animals are well known in the art (*see, e.g.*, Small *et al.*, Mol. Cell. Biol., 5:642-648 [1985]).

G. Animal Models

A variety of well known animal models can be used to further understand the role of the molecules identified herein in the development and pathogenesis of tumors, and to test the efficacy of candidate therapeutic agents, including antibodies, and other agonists of the native polypeptides, including small molecule agonists. The *in vivo* nature of such models makes them particularly predictive of responses in human patients. Animal models of tumors and cancers (*e.g.*, breast cancer, colon cancer, prostate cancer, lung cancer, etc.) include both non-recombinant and recombinant (transgenic) animals. Non-recombinant animal models include, for example, rodent, *e.g.*, murine models. Such models can be generated by introducing tumor cells into syngeneic mice using standard techniques, *e.g.*, subcutaneous injection, tail vein injection, spleen implantation, intraperitoneal implantation, implantation under the renal capsule, or orthotopic implantation, *e.g.*, colon cancer cells implanted in colonic tissue. (*See, e.g.*, PCT publication No. WO 97/33551, published September 18, 1997).

Probably the most often used animal species in oncological studies are immunodeficient mice and, in particular, nude mice. The observation that the nude mouse with hypo/aplasia could successfully act as a host for human tumor xenografts has lead to its widespread use for this purpose. The autosomal recessive *nu* gene has been introduced into a very large number of distinct congenic strains of nude mouse, including, for example, ASW, A/He, AKR, BALB/c, B10.LP, C17, C3H, C57BL, C57, CBA, DBA, DDD, I/st, NC, NFR, NFS, NFS/N, NZB, NZC, NZW, P, RIII and SJL. In addition, a wide variety of other animals with inherited immunological defects other than the nude mouse have been bred and used as recipients of tumor xenografts. For further details *see, e.g.*, The Nude Mouse in Oncology Research, E. Boven and B. Winograd, eds., CRC Press, Inc., 1991.

The cells introduced into such animals can be derived from known tumor/cancer cell lines, such as, any of the above-listed tumor cell lines, and, for example, the B104-1-1 cell line (stable NIH-3T3 cell line transfected with the *neu* protooncogene); *ras*-transfected NIH-3T3 cells; Caco-2 (ATCC HTB-37); a

moderately well-differentiated grade II human colon adenocarcinoma cell line, HT-29 (ATCC HTB-38), or from tumors and cancers. Samples of tumor or cancer cells can be obtained from patients undergoing surgery, using standard conditions, involving freezing and storing in liquid nitrogen (Karmali *et al.*, Br. J. Cancer, 48:689-696 [1983]).

5 Tumor cells can be introduced into animals, such as nude mice, by a variety of procedures. The subcutaneous (s.c.) space in mice is very suitable for tumor implantation. Tumors can be transplanted s.c. as solid blocks, as needle biopsies by use of a trochar, or as cell suspensions. For solid block or trochar implantation, tumor tissue fragments of suitable size are introduced into the s.c. space. Cell suspensions are freshly prepared from primary tumors or stable tumor cell lines, and injected subcutaneously. Tumor cells can
10 also be injected as subdermal implants. In this location, the inoculum is deposited between the lower part of the dermal connective tissue and the s.c. tissue. Boven and Winograd (1991), *supra*. Animal models of breast cancer can be generated, for example, by implanting rat neuroblastoma cells (from which the *neu* oncogen was initially isolated), or *neu*-transformed NIH-3T3 cells into nude mice, essentially as described by Drebin *et al.*, Proc. Natl. Acad. Sci. USA, 83:9129-9133 (1986).

15 Similarly, animal models of colon cancer can be generated by passaging colon cancer cells in animals, *e.g.*, nude mice, leading to the appearance of tumors in these animals. An orthotopic transplant model of human colon cancer in nude mice has been described, for example, by Wang *et al.*, Cancer Research, 54:4726-4728 (1994) and Too *et al.*, Cancer Research, 55:681-684 (1995). This model is based on the so-called "METAMOUSETM" sold by AntiCancer, Inc., (San Diego, California).

20 Tumors that arise in animals can be removed and cultured *in vitro*. Cells from the *in vitro* cultures can then be passaged to animals. Such tumors can serve as targets for further testing or drug screening. Alternatively, the tumors resulting from the passage can be isolated and RNA from pre-passage cells and cells isolated after one or more rounds of passage analyzed for differential expression of genes of interest. Such passaging techniques can be performed with any known tumor or cancer cell lines.

25 For example, Meth A, CMS4, CMS5, CMS21, and WEHI-164 are chemically induced fibrosarcomas of BALB/c female mice (DeLeo *et al.*, J. Exp. Med., 146:720 [1977]), which provide a highly controllable model system for studying the anti-tumor activities of various agents (Palladino *et al.*, J. Immunol., 138:4023-4032 [1987]). Briefly, tumor cells are propagated *in vitro* in cell culture. Prior to injection into the animals, the cell lines are washed and suspended in buffer, at a cell density of about 10×10^6 to 10×10^7 cells/ml. The animals are
30 then infected subcutaneously with 10 to 100 μ l of the cell suspension, allowing one to three weeks for a tumor to appear.

35 In addition, the Lewis lung (3LL) carcinoma of mice, which is one of the most thoroughly studied experimental tumors, can be used as an investigational tumor model. Efficacy in this tumor model has been correlated with beneficial effects in the treatment of human patients diagnosed with small cell carcinoma of the lung (SCCL). This tumor can be introduced in normal mice upon injection of tumor fragments from an affected mouse or of cells maintained in culture (Zupi *et al.*, Br. J. Cancer, 41, suppl. 4:309 [1980]), and evidence indicates that tumors can be started from injection of even a single cell and that a very high proportion of infected tumor cells survive. For further information about this tumor model *see*, Zacharski, Haemostasis,

16:300-320 (1986).

One way of evaluating the efficacy of a test compound in an animal model on an implanted tumor is to measure the size of the tumor before and after treatment. Traditionally, the size of implanted tumors has been measured with a slide caliper in two or three dimensions. The measure limited to two dimensions does not accurately reflect the size of the tumor, therefore, it is usually converted into the corresponding volume by using a mathematical formula. However, the measurement of tumor size is very inaccurate. The therapeutic effects of a drug candidate can be better described as treatment-induced growth delay and specific growth delay. Another important variable in the description of tumor growth is the tumor volume doubling time. Computer programs for the calculation and description of tumor growth are also available, such as the program reported by Rygaard and Spang-Thomsen, Proc. 6th Int. Workshop on Immune-Deficient Animals, Wu and Sheng eds., Basel, 1989, 301. It is noted, however, that necrosis and inflammatory responses following treatment may actually result in an increase in tumor size, at least initially. Therefore, these changes need to be carefully monitored, by a combination of a morphometric method and flow cytometric analysis.

Recombinant (transgenic) animal models can be engineered by introducing the coding portion of the genes identified herein into the genome of animals of interest, using standard techniques for producing transgenic animals. Animals that can serve as a target for transgenic manipulation include, without limitation, mice, rats, rabbits, guinea pigs, sheep, goats, pigs, and non-human primates, *e.g.*, baboons, chimpanzees and monkeys. Techniques known in the art to introduce a transgene into such animals include pronucleic microinjection (Hoppe and Wanger, U.S. Patent No. 4,873,191); retrovirus-mediated gene transfer into germ lines (*e.g.*, Van der Putten *et al.*, Proc. Natl. Acad. Sci. USA, 82:6148-615 [1985]); gene targeting in embryonic stem cells (Thompson *et al.*, Cell, 56:313-321 [1989]); electroporation of embryos (Lo, Mol. Cell. Biol., 3:1803-1814 [1983]); sperm-mediated gene transfer (Lavitrano *et al.*, Cell, 57:717-73 [1989]). For review, *see*, for example, U.S. Patent No. 4,736,866.

For the purpose of the present invention, transgenic animals include those that carry the transgene only in part of their cells ("mosaic animals"). The transgene can be integrated either as a single transgene, or in concatamers, *e.g.*, head-to-head or head-to-tail tandems. Selective introduction of a transgene into a particular cell type is also possible by following, for example, the technique of Lasko *et al.*, Proc. Natl. Acad. Sci. USA, 89:6232-636 (1992).

The expression of the transgene in transgenic animals can be monitored by standard techniques. For example, Southern blot analysis or PCR amplification can be used to verify the integration of the transgene. The level of mRNA expression can then be analyzed using techniques such as *in situ* hybridization, Northern blot analysis, PCR, or immunocytochemistry. The animals are further examined for signs of tumor or cancer development.

The efficacy of antibodies specifically binding the polypeptides identified herein and other drug candidates, can be tested also in the treatment of spontaneous animal tumors. A suitable target for such studies is the feline oral squamous cell carcinoma (SCC). Feline oral SCC is a highly invasive, malignant tumor that is the most common oral malignancy of cats, accounting for over 60% of the oral tumors reported in this species. It rarely metastasizes to distant sites, although this low incidence of metastasis may merely be a reflection of

the short survival times for cats with this tumor. These tumors are usually not amenable to surgery, primarily because of the anatomy of the feline oral cavity. At present, there is no effective treatment for this tumor. Prior to entry into the study, each cat undergoes complete clinical examination, biopsy, and is scanned by computed tomography (CT). Cats diagnosed with sublingual oral squamous cell tumors are excluded from the study. The tongue can become paralyzed as a result of such tumor, and even if the treatment kills the tumor, the animals may not be able to feed themselves. Each cat is treated repeatedly, over a longer period of time. Photographs of the tumors will be taken daily during the treatment period, and at each subsequent recheck. After treatment, each cat undergoes another CT scan. CT scans and thoracic radiograms are evaluated every 8 weeks thereafter. The data are evaluated for differences in survival, response and toxicity as compared to control groups. Positive response may require evidence of tumor regression, preferably with improvement of quality of life and/or increased life span.

In addition, other spontaneous animal tumors, such as fibrosarcoma, adenocarcinoma, lymphoma, chondroma, leiomyosarcoma of dogs, cats, and baboons can also be tested. Of these mammary adenocarcinoma in dogs and cats is a preferred model as its appearance and behavior are very similar to those in humans. However, the use of this model is limited by the rare occurrence of this type of tumor in animals.

H. Screening Assays for Drug Candidates

Screening assays for drug candidates are designed to identify compounds that competitively bind or complex with the receptor(s) of the polypeptides identified herein, or otherwise signal through such receptor(s). Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds, including peptides, preferably soluble peptides, (poly)peptide-immunoglobulin fusions, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

In binding assays, the interaction is binding and the complex formed can be isolated or detected in the reaction mixture. In a particular embodiment, a receptor of a polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, *e.g.*, on a microtiter plate, by covalent or non-covalent attachments. Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the polypeptide and drying. Alternatively, an immobilized antibody, *e.g.*, a monoclonal antibody, specific for the polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, *e.g.*, the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, *e.g.*, by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized

component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

If the candidate compound interacts with but does not bind to a particular receptor, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers [Fields and Song, Nature (London), 340:245-246 (1989); Chien *et al.*, Proc. Natl. Acad. Sci. USA, 88:9578-9582 (1991)] as disclosed by Chevray and Nathans [Proc. Natl. Acad. Sci. USA, 89:5789-5793 (1991)]. Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, while the other one functioning as the transcription activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system") takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-*lacZ* reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for β -galactosidase. A complete kit (MATCHMAKER™) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

I. Pharmaceutical Compositions

The polypeptides of the present invention, agonist antibodies specifically binding proteins identified herein, as well as other molecules identified by the screening assays disclosed herein, can be administered for the treatment of tumors, including cancers, in the form of pharmaceutical compositions.

Where antibody fragments are used, the smallest inhibitory fragment which specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable region sequences of an antibody, peptide molecules can be designed which retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology (*see, e.g.*, Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90:7889-7893 [1993]).

The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

Therapeutic formulations of the polypeptides identified herein, or agonists thereof are prepared for storage by mixing the active ingredient having the desired degree of purity with optional pharmaceutically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences, 16th edition, Osol, A. ed. [1980]), in

the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients, or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid and methionine; preservatives (such as octadecyldimethylbenzyl ammonium chloride; hexamethonium chloride; benzalkonium chloride, benzethonium chloride; phenol, butyl or benzyl alcohol; alkyl parabens such as methyl or propyl paraben; catechol; resorcinol; cyclohexanol; 3-pentanol; and m-cresol); low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, histidine, arginine, or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugars such as sucrose, mannitol, trehalose or sorbitol; salt-forming counter-ions such as sodium; metal complexes (*e.g.*, Zn-protein complexes); and/or non-ionic surfactants such as TWEEN™, PLURONICS™ or polyethylene glycol (PEG).

The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise a cytotoxic agent, cytokine or growth inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, 16th edition, Osol, A. ed. (1980).

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes, prior to or following lyophilization and reconstitution.

Therapeutic compositions herein generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOT™ (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated antibodies remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation

mechanism is discovered to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulfhydryl residues, lyophilizing from acidic solutions, controlling moisture content, using appropriate additives, and developing specific polymer matrix compositions.

J. Methods of Treatment

5 It is contemplated that the polypeptides of the present invention and their agonists, including antibodies, peptides, and small molecule agonists, may be used to treat various tumors, *e.g.*, cancers. Exemplary conditions or disorders to be treated include benign or malignant tumors (*e.g.*, renal, liver, kidney, bladder, breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval, thyroid, hepatic carcinomas; sarcomas; glioblastomas; and various head and neck tumors); leukemias and lymphoid malignancies; other disorders such as neuronal, glial, 10 astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal and blastocoelic disorders; and inflammatory, angiogenic and immunologic disorders. The anti-tumor agents of the present invention (including the polypeptides disclosed herein and agonists which mimic their activity, *e.g.*, antibodies, peptides and small organic molecules), are administered to a mammal, preferably a human, in accord with known methods, such as intravenous administration as a bolus or by continuous infusion over a period of time, or by intramuscular, 15 intraperitoneal, intracerebrospinal, intraocular, intraarterial, intralesional, subcutaneous, intraarticular, intrasynovial, intrathecal, oral, topical, or inhalation routes.

Other therapeutic regimens may be combined with the administration of the anti-cancer agents of the instant invention. For example, the patient to be treated with such anti-cancer agents may also receive radiation therapy. Alternatively, or in addition, a chemotherapeutic agent may be administered to the patient. Preparation and dosing schedules for such chemotherapeutic agents may be used according to manufacturers' instructions 20 or as determined empirically by the skilled practitioner. Preparation and dosing schedules for such chemotherapy are also described in Chemotherapy Service, ed., M.C. Perry, Williams & Wilkins, Baltimore, MD (1992). The chemotherapeutic agent may precede, or follow administration of the anti-tumor agent of the present invention, or may be given simultaneously therewith. The anti-cancer agents of the present invention may be combined with an anti-oestrogen compound such as tamoxifen or an anti-progesterone such as 25 onapristone (*see*, EP 616812) in dosages known for such molecules.

It may be desirable to also administer antibodies against tumor associated antigens, such as antibodies which bind to the ErbB2, EGFR, ErbB3, ErbB4, or vascular endothelial factor (VEGF). Alternatively, or in addition, two or more antibodies binding the same or two or more different cancer-associated antigens may be 30 co-administered to the patient. Sometimes, it may be beneficial to also administer one or more cytokines to the patient. In a preferred embodiment, the anti-cancer agents herein are co-administered with a growth inhibitory agent. For example, the growth inhibitory agent may be administered first, followed by the administration of an anti-cancer agent of the present invention. However, simultaneous administration or administration of the anti-cancer agent of the present invention first is also contemplated. Suitable dosages for the growth inhibitory 35 agent are those presently used and may be lowered due to the combined action (synergy) of the growth inhibitory agent and the antibody herein.

For the prevention or treatment of disease, the appropriate dosage of an anti-tumor agent herein will

depend on the type of disease to be treated, as defined above, the severity and course of the disease, whether the agent is administered for preventive or therapeutic purposes, previous therapy, the patient's clinical history and response to the agent, and the discretion of the attending physician. The agent is suitably administered to the patient at one time or over a series of treatments. Animal experiments provide reliable guidance for the determination of effective doses for human therapy. Interspecies scaling of effective doses can be performed following the principles laid down by Mordenti, J. and Chappell, W. "The use of interspecies scaling in toxicokinetics" in Toxicokinetics and New Drug Development, Yacobi *et al.*, eds., Pergamon Press, New York 1989, pp. 42-96.

For example, depending on the type and severity of the disease, about 1 $\mu\text{g/kg}$ to 15 mg/kg (*e.g.*, 0.1-20 mg/kg) of an antitumor agent is an initial candidate dosage for administration to the patient, whether, for example, by one or more separate administrations, or by continuous infusion. A typical daily dosage might range from about 1 $\mu\text{g/kg}$ to 100 mg/kg or more, depending on the factors mentioned above. For repeated administrations over several days or longer, depending on the condition, the treatment is sustained until a desired suppression of disease symptoms occurs. However, other dosage regimens may be useful. The progress of this therapy is easily monitored by conventional techniques and assays. Guidance as to particular dosages and methods of delivery is provided in the literature; *see*, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting one organ or tissue, for example, may necessitate delivery in a manner different from that to another organ or tissue.

K. Articles of Manufacture

In another embodiment of the invention, an article of manufacture containing materials useful for the diagnosis or treatment of the disorders described above is provided. The article of manufacture comprises a container and a label. Suitable containers include, for example, bottles, vials, syringes, and test tubes. The containers may be formed from a variety of materials such as glass or plastic. The container holds a composition which is effective for diagnosing or treating the condition and may have a sterile access port (for example the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). The active agent in the composition is an anti-tumor agent of the present invention. The label on, or associated with, the container indicates that the composition is used for diagnosing or treating the condition of choice. The article of manufacture may further comprise a second container comprising a pharmaceutically-acceptable buffer, such as phosphate-buffered saline, Ringer's solution and dextrose solution. It may further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, needles, syringes, and package inserts with instructions for use.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas, VA.

EXAMPLE 1

Extracellular Domain Homology Screening to Identify Novel Polypeptides and cDNA Encoding Therefor

The extracellular domain (ECD) sequences (including the secretion signal sequence, if any) from about 950 known secreted proteins from the Swiss-Prot public database were used to search EST databases. The EST databases included public databases (*e.g.*, Dayhoff, GenBank), and proprietary databases (*e.g.* LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST-2 (Altschul *et al.*, Methods in Enzymology, 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. Those comparisons with a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, WA).

Using this extracellular domain homology screen, consensus DNA sequences were assembled relative to the other identified EST sequences using phrap. In addition, the consensus DNA sequences obtained were often (but not always) extended using repeated cycles of BLAST or BLAST-2 and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above.

Based upon the consensus sequences obtained as described above, oligonucleotides were then synthesized and used to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for a PRO polypeptide. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5 kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel *et al.*, Current Protocols in Molecular Biology, *supra*, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; *see*, Holmes *et al.*, Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

EXAMPLE 2Isolation of cDNA Clones Using Signal Algorithm Analysis

Various polypeptide-encoding nucleic acid sequences were identified by applying a proprietary signal sequence finding algorithm developed by Genentech, Inc., (South San Francisco, CA) upon ESTs as well as clustered and assembled EST fragments from public (*e.g.*, GenBank) and/or private (LIFESEQ®, Incyte Pharmaceuticals, Inc., Palo Alto, CA) databases. The signal sequence algorithm computes a secretion signal score based on the character of the DNA nucleotides surrounding the first and optionally the second methionine codon(s) (ATG) at the 5'-end of the sequence or sequence fragment under consideration. The nucleotides following the first ATG must code for at least 35 unambiguous amino acids without any stop codons. If the first ATG has the required amino acids, the second is not examined. If neither meets the requirement, the candidate sequence is not scored. In order to determine whether the EST sequence contains an authentic signal sequence, the DNA and corresponding amino acid sequences surrounding the ATG codon are scored using a set of seven sensors (evaluation parameters) known to be associated with secretion signals. Use of this algorithm resulted in the identification of numerous polypeptide-encoding nucleic acid sequences.

EXAMPLE 3Isolation of cDNA Clones Encoding Human PRO943

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in EXAMPLE 1 above and is herein designated DNA36360. In some cases, the DNA36360 consensus sequence derives from an intermediate consensus DNA sequence which was extended using repeated cycles of BLAST and phrap to extend that intermediate consensus sequence as far as possible using the sources of EST sequences discussed above.

Based on the DNA36360 consensus sequence, oligonucleotides were synthesized : 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO943. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5 kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel *et al.*, Current Protocols in Molecular Biology, *supra*, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CGAGATGACGCCGAGCCCC-3' (SEQ ID NO:7)

reverse PCR primer 5'-CGGTTCGACACGCGGCAGGTG-3' (SEQ ID NO:8)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA36360 sequence which had the following nucleotide sequence:

hybridization probe

5'-TGCTGCTCCTGCTGCCGCCGCTGCTGCTGGGGGCCTTCCCGCCGG-3' (SEQ ID NO:9)

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue. The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; *see*, Holmes *et al.*, Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for a full-length PRO943 polypeptide (designated herein as DNA52192-1369 [Figure 1, SEQ ID NO:1]) and the derived protein sequence for that PRO943 polypeptide.

The full length clone identified above contained a single open reading frame with an apparent translational initiation site at nucleotide positions 150-152 and a stop signal at nucleotide positions 1662-1664 (Figure 1, SEQ ID NO:1). The predicted polypeptide precursor is 504 amino acids long (Figure 2), has a calculated molecular weight of approximately 54,537 daltons and an estimated pI of approximately 10.04. Analysis of the full length PRO943 sequence shown in Figure 2 (SEQ ID NO:2) evidences the presence of a variety of important polypeptide domains as shown in Figure 2, wherein the locations given for those important polypeptide domains are approximate as described above. Analysis of the full-length PRO943 sequence shown in Figure 2 evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 17; a transmembrane domain from about amino acid 376 to about amino acid 396; tyrosine kinase phosphorylation sites from about amino acid 212 to about amino acid 220 and from about amino acid 329 to about amino acid 337; potential N-glycosylation sites from about amino acid 111 to about amino acid 115, from about amino acid 231 to about amino acid 235, from about amino acid 255 to about amino acid 259, and from about amino acid 293 to about amino acid 297; and an immunoglobulin and MHC protein sequence homology block from about amino acid 219 to about amino acid 237. Clone DNA52192-1369 has been deposited with the ATCC on July 1, 1998 and is assigned ATCC deposit no. 203042.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using the ALIGN-2 sequence alignment analysis of the full-length sequence shown in Figure 2 (SEQ ID NO:2) evidenced sequence identity between the PRO943 amino acid sequence and the following Dayhoff sequences: B49151, A39752, FGR1_XENLA, S38579, RATHBFGFRB_1, TVHU2F, FGR2_MOUSE, CEK3_CHICK, P_R21080 and A27171_1.

EXAMPLE 4

Isolation of cDNA clones Encoding Human PRO1250

Use of the signal algorithm procedure described above in EXAMPLE 2 resulted in the identification of an EST cluster sequence from the Incyte database, designated Incyte EST cluster sequence no. 56523. This sequence was then compared to a variety of various EST databases as described under the signal algorithm procedure above, and further resulted in the identification of Incyte EST 3371784. Further examination and sequencing of the full-length clone corresponding to this EST sequence resulted in the isolation of the full-length DNA sequence DNA60775-1532 (Figure 3, SEQ ID NO:3) and the derived PRO1250 native

sequence protein (Figure 4, SEQ ID NO:4).

Clone DNA60775-1532 (SEQ ID NO:3) contains a single open reading frame with an apparent translation initiation site at nucleotide positions 74-76 and ending at the stop codon (TAG) at nucleotide positions 2291-2293 (Figure 3). The predicted PRO1250 polypeptide precursor (SEQ ID NO:4) is 739 amino acids long (Figure 4). The PRO1250 protein shown in Figure 4 has an estimated molecular weight of about 82,263 daltons and a pI of about 7.55. Analysis of PRO1250 polypeptide (SEQ ID NO:4) evidences the presence of the following: a type II transmembrane domain from about amino acid residues 61 to about 80, a putative AMP-binding domain signature sequence from about amino acid residue 314 to about 326, and potential N-glycosylation sites from about amino acid residues 102 to about 106, from about amino acid residues 588 to about 594 and from about amino acid residues 619 to about 623. A cDNA clone containing DNA60775-1250 (SEQ ID NO:3) has been deposited with the ATCC on September 1, 1998 and is assigned ATCC deposit no. 203173.

EXAMPLE 5

Isolation of cDNA clones Encoding Human PRO1337

Using the method described in Example 1 above, a single Incyte EST was identified (EST No.1747546) and also referred to herein as "DNA4417". To assemble a consensus sequence, repeated cycles of BLAST and phrap were used to extend the DNA4417 sequence as far as possible using the sources of EST sequences discussed above. The consensus sequence is designated herein as "DNA45669". Based on the DNA45669 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1337.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 45699.f1:

5'-CAACCATGCAAGGACAGGGCAGG-3' (SEQ ID NO:10)

forward PCR primer 45699.f2:

5'-CTTTGCTGTTGGCCTCTGTGCTCCCAACCATGCAAGGACAGGGCAGG-3' (SEQ ID NO:11)

reverse PCR primer 45669.r1:

5'-TGACTCGGGGTCTCCAAAACCAGC-3' (SEQ ID NO:12)

reverse PCR primer 45669.r2:

5'-GGTATAGGCGGAAGGCAAAGTCGG-3' (SEQ ID NO:13)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA45669 sequence which had the following nucleotide sequence:

hybridization probe 45669.p1:

5'-GGCATCTTACCTTTATGGAGTACTCTTTGCTGTTGGCCTCTGTGCTCC-3' (SEQ ID NO:14)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1337 gene using the probe oligonucleotide and one of the PCR primers. RNA for

construction of the cDNA libraries was isolated from human tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1337 (designated herein as DNA66672-1586 [Figure 5, SEQ ID NO:5]; and the derived protein sequence for PRO1337.

5 The entire coding sequence of PRO1337 is shown in Figure 5 (SEQ ID NO:5). Clone DNA66672-1586 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 60-62 and an apparent stop codon at nucleotide positions 1311-1313. The predicted polypeptide precursor is 417 amino acids long. The full-length PRO1337 protein shown in Figure 6 has an estimated molecular weight of about 46,493 daltons and a pI of about 9.79.

10 Analysis of PRO1337 polypeptide (SEQ ID NO:6) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 20; an N-glycosylation sites from about amino acid 101 to about amino acid 105 and from about amino acid 390 to about amino acid 394; a tyrosine kinase phosphorylation site from about amino acid 377 to about amino acid 385; and an N-myristoylation sites from about amino acid 7 to about amino acid 13, from about amino acid 97 to about amino acid 103, from about amino acid 326 to about amino acid 332, and from about amino acid 363 to about amino acid 369.

15 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 6 (SEQ ID NO:6) revealed significant homology between the PRO1337 amino acid sequence and the Dayhoff sequence THBG_HUMAN. Homology was also found between the PRO1337 amino acid sequence and the following Dayhoff sequences: KAIN_HUMAN, HSACT1_1, IPSP_HUMAN, G02081, HAMHPP_1, CPI6_RAT, S31507, AB000547_1, and KBP_MOUSE.

20 Clone DNA66672-1586 was deposited with the ATCC on September 22, 1998, and is assigned ATCC deposit no. 203265.

EXAMPLE 6

In situ Hybridization

25 *In situ* hybridization is a powerful and versatile technique for the detection and localization of nucleic acid sequences within cell or tissue preparations. It may be useful, for example, to identify sites of gene expression, analyze the tissue distribution of transcription, identify and localize viral infection, follow changes in specific mRNA synthesis, and aid in chromosome mapping.

30 *In situ* hybridization was performed following an optimized version of the protocol by Lu and Gillett, Cell Vision, 1: 169-176 (1994), using PCR-generated ³³P-labeled riboprobes. Briefly, formalin-fixed, paraffin-embedded human tissues were sectioned, deparaffinized, deproteinized in proteinase K (20 g/ml) for 15 minutes at 37°C, and further processed for *in situ* hybridization as described by Lu and Gillett, *supra*. A (³³-P)UTP-labeled antisense riboprobe was generated from a PCR product and hybridized at 55°C overnight. The slides were dipped in Kodak NTB2™ nuclear track emulsion and exposed for 4 weeks.

³³P-Riboprobe synthesis

35 6.0 μl (125 mCi) of ³³P-UTP (Amersham BF 1002, SA<2000 Ci/mmol) were speed-vacuum dried. To each tube containing dried ³³P-UTP, the following ingredients were added:

- 2.0 μ l 5x transcription buffer
 1.0 μ l DTT (100 mM)
 2.0 μ l NTP mix (2.5 mM: 10 μ l each of 10 mM GTP, CTP & ATP + 10 μ l H₂O)
 1.0 μ l UTP (50 μ M)
 5 1.0 μ l RNAsin
 1.0 μ l DNA template (1 μ g)
 1.0 μ l H₂O
 1.0 μ l RNA polymerase (for PCR products T3 = AS, T7 = S, usually)

10 The tubes were incubated at 37°C for one hour. A total of 1.0 μ l RQ1 DNase was added, followed by incubation at 37°C for 15 minutes. A total of 90 μ l TE (10 mM Tris pH 7.6/1 mM EDTA pH 8.0) was added, and the mixture was pipetted onto DE81 paper. The remaining solution was loaded in a MICROCON-50™ ultrafiltration unit, and spun using program 10 (6 minutes). The filtration unit was inverted over a second tube and spun using program 2 (3 minutes). After the final recovery spin, a total of 100 μ l TE was added, then 1 μ l of the final product was pipetted on DE81 paper and counted in 6 ml of BIOFLUOR II™.

15 The probe was run on a TBE/urea gel. A total of 1-3 μ l of the probe or 5 μ l of RNA Mrk III was added to 3 μ l of loading buffer. After heating on a 95°C heat block for three minutes, the gel was immediately placed on ice. The wells of gel were flushed, and the sample was loaded and run at 180-250 volts for 45 minutes. The gel was wrapped in plastic wrap (SARAN™ brand) and exposed to XAR film with an intensifying screen in a -70°C freezer one hour to overnight.

20 ³³P-Hybridization

A. *Pretreatment of frozen sections*

The slides were removed from the freezer, placed on aluminum trays, and thawed at room temperature for 5 minutes. The trays were placed in a 55°C incubator for five minutes to reduce condensation. The slides were fixed for 10 minutes in 4% paraformaldehyde on ice in the fume hood, and washed in 0.5 x SSC for 5 minutes, at room temperature (25 ml 20 x SSC + 975 ml SQ H₂O). After deproteinization in 0.5 μ g/ml proteinase K for 10 minutes at 37°C (12.5 μ l of 10 mg/ml stock in 250 ml prewarmed RNase-free RNase buffer), the sections were washed in 0.5 x SSC for 10 minutes at room temperature. The sections were dehydrated in 70%, 95%, and 100% ethanol, 2 minutes each.

B. *Pretreatment of paraffin-embedded sections*

30 The slides were deparaffinized, placed in SQ H₂O, and rinsed twice in 2 x SSC at room temperature, for 5 minutes each time. The sections were deproteinized in 20 μ g/ml proteinase K (500 μ l of 10 mg/ml in 250 ml RNase-free RNase buffer; 37°C, 15 minutes) for human embryo tissue, or 8 x proteinase K (100 μ l in 250 ml RNase buffer, 37°C, 30 minutes) for formalin tissues. Subsequent rinsing in 0.5 x SSC and dehydration were performed as described above.

35 C. *Prehybridization*

The slides were laid out in a plastic box lined with Box buffer (4 x SSC, 50% formamide) - saturated filter paper. The tissue was covered with 50 μ l of hybridization buffer (3.75 g dextran sulfate + 6 ml SQ H₂O),

vortexed, and heated in the microwave for 2 minutes with the cap loosened. After cooling on ice, 18.75 ml formamide, 3.75 ml 20 x SSC, and 9 ml SQ H₂O were added, and the tissue was vortexed well and incubated at 42°C for 1-4 hours.

D. *Hybridization*

5 1.0 x 10⁶ cpm probe and 1.0 µl tRNA (50 mg/ml stock) per slide were heated at 95°C for 3 minutes. The slides were cooled on ice, and 48 µl hybridization buffer was added per slide. After vortexing, 50 µl ³³P mix was added to 50 µl prehybridization on the slide. The slides were incubated overnight at 55°C.

E. *Washes*

10 Washing was done for 2x10 minutes with 2xSSC, EDTA at room temperature (400 ml 20 x SSC + 16 ml 0.25 M EDTA, V_f=4L), followed by RNaseA treatment at 37°C for 30 minutes (500 µl of 10 mg/ml in 250 ml Rnase buffer = 20 µg/ml). The slides were washed 2 x10 minutes with 2 x SSC, EDTA at room temperature. The stringency wash conditions were as follows: 2 hours at 55°C, 0.1 x SSC, EDTA (20 ml 20 x SSC + 16 ml EDTA, V_f=4L).

F. *Oligonucleotides*

15 *In situ* analysis was performed on one of the DNA sequences disclosed herein. The oligonucleotides employed for these analyses are as follows:

(1) DNA52192-1369

D-269E FGFr p1:

5'-GGATTCTAATACGACTCACTATAGGGCCGCTGACCATGTGGACCAAGG-3' (SEQ ID NO:15)

20 C-257I FGFr p2:

5'-CTATGAAATTAACCCTCACTAAAGGGATCTGGCAGCACGGTGAGGAAG-3' (SEQ ID NO:16)

G. *Results*

In situ analysis was performed on the above DNA sequence disclosed herein. The results from this analysis are as follows:

25 DNA52192-1369 (FGF receptor-like molecule)

30 Expression was observed in fetal skeletal muscle and long bone cartilage. Elsewhere, a relatively high background signal is a problem. In one fetal liver, expression appears to occur at sites of hematopoiesis. The other fetal liver was negative. Fetal tissues examined (E12-E16 weeks) included: placenta, umbilical cord, liver, kidney, adrenals, thyroid, lung, heart, great vessels, esophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb. Adult human tissues examined included: liver, kidney, adrenals, myocardium, aorta, spleen, lung, skin, chondrosarcoma, eye, stomach, gastric carcinoma, colon, colonic carcinoma, renal cell carcinoma, prostate, bladder mucosa and gall bladder. Acetaminophen induced liver injury and hepatic cirrhosis. Rhesus tissues examined included cerebral cortex (rm) and hippocampus (rm). Chimp tissues examined included: thyroid, parathyroid, ovary, nerve, tongue, thymus, 35 adrenals, gastric mucosa and salivary gland.

EXAMPLE 7

Use of PRO as a Hybridization Probe

The following method describes use of a nucleotide sequence encoding PRO as a hybridization probe.

DNA comprising the coding sequence of full-length or mature PRO as disclosed herein or a fragment thereof is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high-stringency conditions. Hybridization of radiolabeled probe derived from the gene encoding a PRO polypeptide to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO can then be identified using standard techniques known in the art.

EXAMPLE 8

Expression of PRO in *E. coli*

This example illustrates preparation of an unglycosylated form of PRO by recombinant expression in *E. coli*.

The DNA sequence encoding PRO is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see, Bolivar *et al.*, Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a poly-His leader (including the first six STII codons, poly-His sequence, and enterokinase cleavage site), the PRO coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook *et al.*, *supra*. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

PRO may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO is initially amplified using selected PCR primers. The primers will contain restriction enzyme

sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an OD₆₀₀ of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g (NH₄)₂SO₄, 0.71 g sodium citrate•2H₂O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 ml water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM MgSO₄) and grown for approximately 20-30 hours at 30°C with shaking. Samples are removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

E. coli paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentrifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The clarified extract is loaded onto a 5 ml Qiagen Ni²⁺-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4°C for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A₂₈₀ absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

Fractions containing the desired folded PRO polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with

0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 9

Expression of PRO in mammalian cells

This example illustrates preparation of a potentially glycosylated form of PRO by recombinant expression in mammalian cells.

The vector, pRK5 (*see* EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO DNA using ligation methods such as described in Sambrook *et al.*, *supra*. The resulting vector is called pRK5-PRO.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 μ g pRK5-PRO DNA is mixed with about 1 μ g DNA encoding the VA RNA gene [Thimmappaya *et al.*, *Cell*, 31:543 (1982)] and dissolved in 500 μ l of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl_2 . To this mixture is added, dropwise, 500 μ l of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO_4 , and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μ Ci/ml ^{35}S -cysteine and 200 μ Ci/ml ^{35}S -methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of the PRO polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac *et al.*, *Proc. Natl. Acad. Sci.*, 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μ g pRK5-PRO DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μ g/ml bovine insulin and 0.1 μ g/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO can then be concentrated and purified by any selected

method, such as dialysis and/or column chromatography.

In another embodiment, PRO can be expressed in CHO cells. The pRK5-PRO can be transfected into CHO cells using known reagents such as CaPO_4 or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ^{35}S -methionine. After determining the presence of a PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO polypeptide can then be concentrated and purified by any selected method.

Epitope-tagged PRO may also be expressed in host CHO cells. The PRO may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-His tag into a Baculovirus expression vector. The poly-His tagged PRO insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO can then be concentrated and purified by any selected method, such as by Ni^{2+} -chelate affinity chromatography.

PRO may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (*e.g.*, extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or as a poly-His tagged form.

Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using standard techniques as described in Ausubel *et al.*, Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used in expression in CHO cells is as described in Lucas *et al.*, Nucl. Acids Res., 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect® (Qiagen), Dospers® or Eugene® (Boehringer Mannheim). The cells are grown as described in Lucas *et al.*, *supra*. Approximately 3×10^7 cells are frozen in an ampule for further growth and production as described below.

The ampules containing the plasmid DNA are thawed by placement into a water bath and mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 ml of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 ml of selective media (0.2 μm filtered PS20 with 5% 0.2 μm diafiltered fetal bovine serum). The cells are then aliquoted into a 100 ml spinner containing 90 ml of selective media. After 1-2 days, the cells are transferred into a 250 ml spinner filled with 150 ml selective growth medium and incubated at 37°C. After another 2-3 days, 250 ml, 500 ml and

2000 ml spinners are seeded with 3×10^5 cells/ml. The cell media is exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Patent No. 5,122,469, issued June 16, 1992 may actually be used. A 3L production spinner is seeded at 1.2×10^6 cells/ml. On day 0, the cell number and pH is determined. On day 1, the spinner is sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, the temperature shifted to 33°C, and 30 ml of 500 g/L glucose and 0.6 ml of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability drops below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22 μ m filter. The filtrate is either stored at 4°C or immediately loaded onto columns for purification.

For the poly-His tagged constructs, the proteins are purified using a Ni²⁺-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni²⁺-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The conditioned medium is pumped onto a 5 ml Protein A column (Pharmacia) which has been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275 μ l of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 10

Expression of PRO in Yeast

The following method describes recombinant expression of PRO in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO from the ADH2/GAPDH promoter. DNA encoding PRO and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of PRO. For secretion, DNA encoding PRO can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native PRO signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with

Coomassie Blue stain.

Recombinant PRO can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO may further be purified using selected column chromatography resins.

5 Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 11

Expression of PRO in Baculovirus-Infected Insect Cells

The following method describes recombinant expression in Baculovirus-infected insect cells.

10 The sequence coding for PRO is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-His tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding PRO or the desired portion of the coding sequence of PRO (such as the sequence encoding the extracellular domain of a transmembrane protein or the sequence encoding the mature protein if the protein is extracellular) is amplified by PCR with primers complementary to 15 the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and 20 used for further amplifications. Viral infection and protein expression are performed as described by O'Reilley *et al.*, Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-His tagged PRO can then be purified, for example, by Ni²⁺-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert *et al.*, Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 ml Hepes, pH 7.9; 12.5 mM MgCl₂; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 25 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 µm filter. A Ni²⁺-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 ml, washed with 25 ml of water and equilibrated with 25 ml of loading buffer. The filtered cell extract is loaded onto the column at 0.5 ml per minute. The column is washed to baseline A₂₈₀ with loading buffer, at which point fraction 30 collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A₂₈₀ baseline again, the column is developed with a 0 to 500 mM imidazole gradient in the secondary wash buffer. One ml fractions are collected and analyzed by SDS-PAGE and silver staining or Western blot with Ni²⁺-NTA-conjugated to 35 alkaline phosphatase (Qiagen). Fractions containing the eluted His₁₀-tagged PRO are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO can be performed using known

chromatography techniques, including for instance, Protein A or protein G column chromatography.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 12

Preparation of Antibodies that Bind PRO

5 This example illustrates preparation of monoclonal antibodies which can specifically bind the PRO polypeptide or an epitope on the PRO polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, *supra*. Immunogens that may be employed include purified PRO, fusion proteins containing PRO,
10 and cells expressing recombinant PRO on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected
15 into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with
20 a final intravenous injection of PRO. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

25 The hybridoma cells will be screened in an ELISA for reactivity against PRO. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against PRO is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be
30 accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

EXAMPLE 13

Purification of PRO Polypeptides Using Specific Antibodies

Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the art of
35 protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In

general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such an immunoaffinity column is utilized in the purification of the PRO polypeptide by preparing a fraction from cells containing the PRO polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of the PRO polypeptide (*e.g.*, high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO polypeptide binding (*e.g.*, a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and the PRO polypeptide is collected.

EXAMPLE 14

Drug Screening

This invention is particularly useful for screening compounds by using PRO polypeptides or a binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between a PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with a PRO polypeptide or fragment thereof and assaying (i) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or fragment is typically labeled. After suitable incubation, the free PRO polypeptide or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the

ability of the particular agent to bind to the PRO polypeptide or to interfere with the PRO polypeptide/cell complex.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with the PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding a PRO polypeptide specifically compete with a test compound for binding to the PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with a PRO polypeptide.

EXAMPLE 15

Rational Drug Design

The goal of rational drug design is to produce structural analogs of a biologically active polypeptide of interest (*i.e.*, a PRO polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide *in vivo* (*c.f.*, Hodgson, Bio/Technology, 9: 19-21 (1991)).

In one approach, the three-dimensional structure of the PRO polypeptide, or of a PRO polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, Biochemistry, 31:7796-7801 (1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, J. Biochem., 113:742-746 (1993).

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The

isolated peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

EXAMPLE 16

In Vitro Antitumor Assay

The antiproliferative activity of the PRO943, PRO1250, and PRO1337 polypeptide was determined in the investigational, disease-oriented *in vitro* anti-cancer drug discovery assay of the National Cancer Institute (NCI), using a sulforhodamine B (SRB) dye binding assay essentially as described by Skehan *et al.*, J. Natl. Cancer Inst., 82:1107-1112 (1990). The 60 tumor cell lines employed in this study ("the NCI panel"), as well as conditions for their maintenance and culture *in vitro* have been described by Monks *et al.*, J. Natl. Cancer Inst., 83:757-766 (1991). The purpose of this screen is to initially evaluate the cytotoxic and/or cytostatic activity of the test compounds against different types of tumors (Monks *et al.*, *supra*; Boyd, Cancer: Princ. Pract. Oncol. Update, 3(10):1-12 [1989]).

Cells from approximately 60 human tumor cell lines were harvested with trypsin/EDTA (Gibco), washed once, resuspended in IMEM and their viability was determined. The cell suspensions were added by pipet (100 μ l volume) into separate 96-well microtiter plates. The cell density for the 6-day incubation was less than for the 2-day incubation to prevent overgrowth. Inoculates were allowed a preincubation period of 24 hours at 37°C for stabilization. Dilutions at twice the intended test concentration were added at time zero in 100 μ l aliquots to the microtiter plate wells (1:2 dilution). Test compounds were evaluated at five half-log dilutions (1000 to 100,000-fold). Incubations took place for two days and six days in a 5% CO₂ atmosphere and 100% humidity.

After incubation, the medium was removed and the cells were fixed in 0.1 ml of 10% trichloroacetic acid at 40°C. The plates were rinsed five times with deionized water, dried, stained for 30 minutes with 0.1 ml of 0.4% sulforhodamine B dye (Sigma) dissolved in 1% acetic acid, rinsed four times with 1% acetic acid to remove unbound dye, dried, and the stain was extracted for five minutes with 0.1 ml of 10 mM Tris base [tris(hydroxymethyl)aminomethane], pH 10.5. The absorbance (OD) of sulforhodamine B at 492 nm was measured using a computer-interfaced, 96-well microtiter plate reader.

A test sample is considered positive if it shows at least 40% growth inhibitory effect at one or more concentrations. The results are shown in the following Table 7, where the tumor cell type abbreviations are as follows:

NSCL = non-small cell lung carcinoma; CNS = central nervous system

Table 7

	<u>Test compound</u>	<u>Days</u>	<u>Tumor Cell Line Type</u>	<u>Cell Line Designation</u>
	PRO943	N/A	NSCL	HOP-92, NCI-H522
	PRO943	N/A	Colon	KM12
5	PRO943	N/A	Breast	HS578T
	PRO943	N/A	Ovarian	OVCAR-3
	PRO943	N/A	Leukemia	CCRF-CEM, HL-60(TB), MOLT-4, RPMI-8226
	PRO943	N/A	Melanoma	LOXIMVI
10	PRO943	N/A	CNS	SNB-19
	PRO1250	N/A	NSCL	NCI-H23, NCI-H522
	PRO1250	N/A	CNS	SF-268, SNB-19, U251
	PRO1337	N/A	Leukemia	RPMI-8226

15 Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209, USA (ATCC):

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA52192-1369	203042	July 1, 1998
20	DNA60775-1532	203173	September 1, 1998
	DNA66672-1586	203265	September 22, 1998

25 These deposits were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc., and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 U.S.C. § 122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14 with particular reference to 886 OG 638).

30 The assignee of the present application has agreed that if a culture of the materials on deposit should

die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

- 5 The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any
- 10 aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

PCT

Original (for SUBMISSION) - printed on 08.11.2000 02:18:25 PM

P3134R1

0-1	Form - PCT/RO/134 (EASY) Indications Relating to Deposited Microorganism(s) or Other Biological Material (PCT Rule 13bis)	
0-1-1	Prepared using	PCT-EASY Version 2.91 (updated 10.10.2000)
0-2	International Application No.	
0-3	Applicant's or agent's file reference	P3134R1
1	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
1-1	page	87
1-2	line	19
1-3	Identification of Deposit	
1-3-1	Name of depositary institution	American Type Culture Collection
1-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
1-3-3	Date of deposit	01 July 1998 (01.07.1998)
1-3-4	Accession Number	ATCC 203042
1-4	Additional Indications	NONE
1-5	Designated States for Which Indications are Made	all designated States
1-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
2	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
2-1	page	87
2-2	line	20
2-3	Identification of Deposit	
2-3-1	Name of depositary institution	American Type Culture Collection
2-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
2-3-3	Date of deposit	01 September 1998 (01.09.1998)
2-3-4	Accession Number	ATCC 203173
2-4	Additional Indications	NONE
2-5	Designated States for Which Indications are Made	all designated States
2-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
3	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
3-1	page	87
3-2	line	21

PCT

P3134R1

Original (for SUBMISSION) - printed on 08.11.2000 02:18:25 PM

3-3	Identification of Deposit	
3-3-1	Name of depositary institution	American Type Culture Collection
3-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
3-3-3	Date of deposit	22 September 1998 (22.09.1998)
3-3-4	Accession Number	ATCC 203265
3-4	Additional Indications	NONE
3-5	Designated States for Which Indications are Made	all designated States
3-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	

FOR RECEIVING OFFICE USE ONLY

0-4	This form was received with the international application: (yes or no)	<i>March 16, 1998</i> Yes
0-4-1	Authorized officer	<i>March 16, 1998</i>

FOR INTERNATIONAL BUREAU USE ONLY

0-5	This form was received by the international Bureau on:	04.01.01
0-5-1	Authorized officer	

WHAT IS CLAIMED IS:

1. A composition of matter useful for the inhibition of neoplastic cell growth, said composition comprising an effective amount of a PRO943, PRO1250, or PRO1337 polypeptide, or an agonist thereof, in admixture with a pharmaceutically acceptable carrier.
2. The composition of matter of Claim 1 comprising a growth inhibitory amount of a PRO943, PRO1250, or PRO1337 polypeptide, or an agonist thereof.
3. The composition of matter of Claim 1 comprising a cytotoxic amount of a PRO943, PRO1250, or PRO1337 polypeptide, or an agonist thereof.
4. The composition of matter of Claim 1 additionally comprising a further growth inhibitory agent, cytotoxic agent or chemotherapeutic agent.
5. A composition of matter useful for the treatment of a tumor in a mammal, said composition comprising a therapeutically effective amount of a PRO943, PRO1250, or PRO1337 polypeptide, or an agonist thereof.
6. The composition of matter of Claim 5, wherein said tumor is a cancer.
7. The composition of matter of Claim 6, wherein the cancer is selected from the group consisting of breast cancer, ovarian cancer, colorectal cancer, lung cancer, central nervous system cancer, melanoma and leukemia.
8. A method for inhibiting the growth of a tumor cell comprising exposing said tumor cell to an effective amount of a PRO943, PRO1250, or PRO1337 polypeptide, or an agonist thereof.
9. The method of Claim 8, wherein said agonist is an anti-PRO943, anti-PRO1250, or anti-PRO1337 agonist antibody.
10. The method of Claim 8, wherein said agonist is a small molecule mimicking the biological activity of a PRO943, PRO1250, or PRO1337 polypeptide.
11. The method of Claim 8, wherein said step of exposing occurs *in vitro*.
12. The method of Claim 8, wherein said step of exposing occurs *in vivo*.
13. An article of manufacture comprising:

a container; and

a composition comprising an active agent contained within the container; wherein said active agent in the composition is a PRO943, PRO1250, or PRO1337 polypeptide, or an agonist thereof.

14. The article of manufacture of Claim 13, further comprising a label affixed to said container, or a package insert included in said container, referring to the use of said composition for the inhibition of neoplastic cell growth.

15. The article of manufacture of Claim 13, wherein said agonist is an anti-PRO943, anti-PRO1250, or anti-PRO1337 agonist antibody.

16. The article of manufacture of Claim 13, wherein said agonist is a small molecule mimicking the biological activity of a PRO943, PRO1250, or PRO1337 polypeptide.

17. The article of manufacture of Claim 13, wherein said active agent is present in an amount that is effective for the treatment of tumor in a mammal.

18. The article of manufacture of Claim 13, wherein said composition additionally comprises a further growth inhibitory agent, cytotoxic agent or chemotherapeutic agent.

19. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence that encodes an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), and Figure 6 (SEQ ID NO:6).

20. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), and Figure 5 (SEQ ID NO:5).

21. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the full-length coding sequence of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), and Figure 5 (SEQ ID NO:5).

22. Isolated nucleic acid having at least 80% nucleic acid sequence identity to the full-length coding sequence of the DNA deposited under ATCC accession number 203042, 203173, or 203265.

23. A vector comprising the nucleic acid of any one of Claims 19 to 22.

24. The vector of Claim 23 operably linked to control sequences recognized by a host cell

transformed with the vector.

25. A host cell comprising the vector of Claim 23.
26. The host cell of Claim 25, wherein said cell is a CHO cell.
27. The host cell of Claim 25, wherein said cell is an *E. coli*.
28. The host cell of Claim 25, wherein said cell is a yeast cell.
29. The host cell of Claim 25, wherein said cell is a Baculovirus-infected insect cell.
30. A process for producing a PRO943, PRO1250, or PRO1337 polypeptide comprising culturing the host cell of Claim 25 under conditions suitable for expression of said polypeptide and recovering said polypeptide from the cell culture.
31. An isolated polypeptide having at least 80% amino acid sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), and Figure 6 (SEQ ID NO:6).
32. An isolated polypeptide having at least 80% amino acid sequence identity to an amino acid sequence encoded by the full-length coding sequence of the DNA deposited under ATCC accession number 203042, 203173, or 203265.
33. A chimeric molecule comprising a polypeptide according to any one of Claims 31 to 32 fused to a heterologous amino acid sequence.
34. The chimeric molecule of Claim 33, wherein said heterologous amino acid sequence is an epitope tag sequence.
35. The chimeric molecule of Claim 33, wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.
36. An isolated antibody which specifically binds to a polypeptide according to any one of Claims 31 to 32.
37. The isolated antibody of Claim 36, wherein said antibody specifically binds to said polypeptide or specifically binds to an epitope on said polypeptide without substantially binding to any other polypeptide

or polypeptide epitope.

38. The isolated antibody of Claim 36 which is a monoclonal antibody, a humanized antibody or a single-chain antibody.

39. Isolated nucleic acid having at least 80% nucleic acid sequence identity to:

- (a) a nucleotide sequence encoding the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), or Figure 6 (SEQ ID NO:6), lacking its associated signal peptide;
- (b) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), or Figure 6 (SEQ ID NO:6), with its associated signal peptide; or
- (c) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), or Figure 6 (SEQ ID NO:6), lacking its associated signal peptide.

40. An isolated polypeptide having at least 80% amino acid sequence identity to:

- (a) the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), or Figure 6 (SEQ ID NO:6), lacking its associated signal peptide;
- (b) an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), or Figure 6 (SEQ ID NO:6), with its associated signal peptide; or
- (c) an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), or Figure 6 (SEQ ID NO:6), lacking its associated signal peptide.

1/6

FIGURE 1

GCCGCCCCGCCCCGAGACCGGGCCCCGGGGGCGCGGGGCGGCGGGATGCGGCGCCCCGGGGCGGCGATGACCG
CGGAGCGCACGCCGCGGGCCCCGCCCCCTGACCCCCGCGCCCCGCTGAGCCCCCGCGAGGTCCGGACA
GGCCGAGATGACGCCGAGCCCCCTGTTGCTGCTCCTGCTGCCGCCGCTGCTGCTGGGGGCTTCCCACCGG
CCGCCCGCGCCCGAGGCCCCCAAGATGGCGGACAAGGTGGTCCACCGGCAGGTGGCCCCGGCTGGGCCGC
ACTGTGCGGCTGCAGTGCCCATGTGGAGGGGACCCGCGCGCTGACCATGTGGACCAAGGATGGCCGCAC
CATCCACAGCGGCTGGAGCCGCTTCCGCGTGCTGCCGCGAGGGGCTGAAGGTGAAGCAGGTGGAGCGGGAGG
ATGCCGCGCTGTACGTGTGCAAGGCCACCAACGGCTTCGGCAGCCTGAGCGTCAACTACACCCCTCGTCGTG
CTGGATGACATTAGCCAGGGAAGGAGAGCCTGGGGCCGACAGCTCCTCTGGGGGTCAAGAGGACCCCGCC
AGCCAGCAGTGGGCACGACCGCGCTTCCACACAGCCCTCCAAGATGAGGCGCCGGGTGATCGCACGGCCCGT
GGGTAGCTCCGTCGGGCTCAAGTGCGTGGCCAGCGGGCACCCCTCGGCCCGACATCACGTGGATGAAGGACG
ACCAGGCCCTTGACGCGCCAGAGGCCGCTGAGCCCAGGAAGAAGTGGACACTGAGCCTGAAGAACCCTG
CGGCCGGAGGACAGCGGCAAATACACCTGCCGCGTGTGCAACCGCGCGGGCGCCATCAACGCCACCTACAA
GGTGGATGTGATCCAGCGGACCCGTTCCAGCTGCAAGGTGCGCAGCGACGTGAAGCCGGTGATCCAGTGGCTGAAG
CGCGTGGAGTACGGCGCCGAGGGCCGCCACAACCTCCACCATCGATGTGGGCGGCCAGAAGTTTGTGGTGCT
GCCACGGGTGACGTGTGGTTCGCGGCCCGACGGCTCCTACCTCAATAAGCTGCTCATACCCGTGCCCGCC
AGGACGATGCGGGCATGTACATCTGCCCTTGGCGCAACACCATGGGCTACAGCTTCCGCGAGCGCCTTCCCTC
ACCGTGTGTCAGACCCAAAACCGCCAGGGCCACCTGTGGCCTCCTCGTCTCGGCCACTAGCCTGCCGTG
GCCCGTGGTATCGGCATCCAGCCGGCGTGTCTTATCCTGGGCACCCCTGCTCCTGTGGCTTTGCCAGG
CCCAGAAGAAGCCGTGCACCCCCGCGCCTGCCCTCCCCCTGCCCTGGGCACCCCGCGGGACGGCCCGC
GACCGCAGCGGAGACAAGGACCTTCCCTCGTTGGCCGCCCTCAGCGCTGGCCCTGGTGTGGGGCTGTGTGA
GGAGCATGGGTCTCCGGCAGCCCCCAGCACTTACTGGGCCAGGCCAGTTGCTGGCCCTAAGTTGTACC
CCAACTCTACACAGACATCCACACACACACACACACTCTCACACACACTCACACGTGGAGGGCAAG
GTCCACCAGCACATCCACTATCAGTGCTAGACGGCACCCGTATCTGCAGTGGGCACGGGGGGGCGGCCAGA
CAGGCAGACTGGGAGGATGGAGGACGGAGCTGCAGACGAAGGCAGGGGACCCATGGCGAGGAGGAATGGCC
AGCACCCCAGGCAGTCTGTGTGTGAGGCATAGCCCCCTGGACACACACACAGACACACACTACCTGGA
TGCATGTATGCACACACATGCGCGCACACGTGCTCCCTGAAGGCACACAGTACGCACACGCACATGCACAGA
TATGCCGCCCTGGGCACACAGATAAGCTGCCCAAATGCACGCACACGCACAGAGACATGCCAGAACATACAA
GGACATGCTGCCTGAACATACACACGCACACCCATGCGCAGATGTGCTGCCCTGGACACACACACACACG
GATATGCTGTCTGGACGCACACACGTGCAGATATGGTATCCGGACACACACGTGCACAGATATGCTGCCCTG
GACACACAGATAATGCTGCCTTGACACACACATGCACGGATATTGCCCTGGACACACACACACACACGCG
TGCACAGATATGCTGTCTGGACACGCACACACATGCAGATATGCTGCCCTGGACACACACTTCCAGACACAC
GTGCACAGGCGCAGATATGCTGCCCTGGACACACGCAGATATGCTGTCTAGTCACACACACACGCAGACATG
CTGTCCGGACACACACACGCATGCACAGATATGCTGTCCGGACACACACACGCACGCAGATATGCTGCCCTG
GACACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGATATTGCCCTGGACACACACATGTGCACAG
ATATGCTGTCTGGACATGCACACACGTGCAGATATGCTGTCCGGATACACACGCACGCACACATGCAGATA
TGCTGCCCTGGGCACACACTTCCGGACACACATGCACACACAGGTGCAGATATGCTGCCCTGGACACACACAC
AGATAATGCTGCCCTCAACACTCACACACGTGCAGATATTGCCCTGGACACACACATGTGCACAGATATGCTG
TCTGGACATGCACACACGTGCAGATATGCTGTCCGGATACACACGCACGCACACATGCAGATATGCTGCCCT
GGGCACACACTTCCGGACACACATGCACACACAGGTGCAGATATGCTGCCCTGGACACACGCAGACTGACGT
GCTTTTGGGAGGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCATAGTTGATGAGGGACTTTC
CCTGCTCCACCGTCACTCCCCAACTCTGCCCGCCTCTGTCCCCGCTCAGTCCCCGCCTCCATCCCCGCC
TCTGTCCCCCTGGCCTTGGCGGCTATTTTGGCCACCTGCCCTTGGGTGCCAGGAGTCCCCTACTGCTGTGGG
CTGGGGTTGGGGGCACAGCAGCCCCAAGCCTGAGAGGCTGGAGCCCATGGCTAGTGCTCATCCCCAGTGC
ATTCTCCCCCTGACACAGAGAAGGGGCCCTTGGTATTTATATTTAAGAAATGAAGATAATATTAATAATGAT
GGAAGGAAGACTGGGTTGCAGGGACTGTGTTCTCCTGGGGCCCGGACCCGCTGGTCTTTCAGCCATG
CTGATGACCACACCCCGTCCAGGCCAGACACACCCCGCTGTCGTGGTGGCCCCAGATCTCTGT
AATTTTATGTAGAGTTTGAGCTGAAGCCCCGTATATTTAATTTATTTTGTAAACACAAAA

2/6

FIGURE 2

MTSPSLLLLLLPPLLLGAFPPAAAAARGPPKMADKVVPRQVARLGRTVRLQCPVEGDPPPLTMWTKDGRTIH
SGWSRFRVLPQGLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVVLDLISPAGESLGPSSSSGGQEDPASQ
QWARPRFTQPSKMRRRVIAIPVGSSVRLKCVASGHPRPDITWMKDDQALTRPEAAEPRKKKWTLNLRP
EDSGKYTCRVSNRAGAINATYKVDVIQRTSRKPVLGTGHPVNTTVDFGGTTSFQCKVRSDVKPVIQWLKRV
EYGAEGRHNSTIDVGGQKFVVLPTGDIVSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRSAFLTV
LPDPKPPGPPVASSSSATSLPWPVVIGIPAGAVFILGTLWWLCQAQKKPCTPAPAPPLPGHRPPGTARDR
SGDKDLPSLAALSAGPGVGLCEEHGSPAAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTSHTHSHVEGKVH
QHIHYQC

Signal peptide:	Amino acids 1-17
Putative transmembrane domain:	Amino acids 376-396
Tyrosine kinase phosphorylation sites:	Amino acids 212-220;329-337
N-glycosylation sites:	Amino acids 111-115;231-235;255-259; 293-297
Immunoglobulins and major histocompatibility protein:	Amino acids 219-237
cAMP- and cGMP-dependent protein kinase phosphorylation site:	Amino acids 202-206
N-myristoylation sites:	Amino acids 95-101;228-234;261-267; 317-323;334-340;382-388;443-449

3/6

FIGURE 3

CTGACATGGCCTGACTCGGGACAGCTCAGAGCAGGGCAGAACTGGGGACACTCTGGGCCGGCCTTCTGCCTGC
ATGGACGCTCTGAAGCCACCCTGTCTCTGGAGGAACCACGAGCGAGGGAAGAAGGACAGGGACTCGTGTGG
CAGGAAGAACTCAGAGCCGGGAAGCCCCCATTCACTAGAAAGCACTGAGAGATGCGGCCCTTCGCAGGGTC
TGAATTTCCCTGCTGCTGTTTCAAAAGATGCTTTTATCTTTAACTTTTGTCTTCCCACTTCCGACCCCG
GCGTTGATCTGCATCCTGACATTGAGAGCTGCCATCTTCTGTGGCTGATCACCAGACCTCAACCCGCTT
ACCTCTTCTTGACCTGAACAATCAGTCTGTGGGAATTGAGGGAGGAGCACGGAAGGGGGTTTCCCAAGAAGA
ACAATGACCTAACAAGTTGCTGCTTCTCAGATGCCAAGACTATGTATGAGGTTTTTCAAAGAGGACTCGCT
GTGTCTGACAATGGGCCCTGCTTGGGATATAGAAAACCAAACAGCCCTACAGATGGCTATCTTACAAACA
GGTGTCTGATAGAGCAGAGTACCTGGGTTCTGTCTCTTGCATAAAGGTTATAAATCATCACCAGACCAGT
TTGTGCGCATCTTTGCTCAGAAATAGGCCAGAGTGGATCATCTCCGAATTGGCTTGTACACGTACTCTATG
GTAGCTGTACCTCTGTATGACACCTTGGGACCAGAAGCCATCGTACATATTGTCAACAAGGCTGATATCGC
CATGGTGATCTGTGACACACCCCAAAAGGCATTGGTGCTGATAGGGAATGTAGAGAAAGGCTTCAACCCGA
GCCTGAAGGTGATCATCTTATGGACCCCTTTGATGATGACCTGAAGCAAAGAGGGGAGAAGAGTGGAATT
GAGATCTTATCCCTATATGATGCTGAGAACCTAGGCAAAGAGCACTTCAGAAAACCTGTGCCTCCTAGCCC
AGAAGACCTGAGCGTCATCTGCTTCAACAGTGGGACCACAGGTGACCCCAAAGGAGCCATGATAACCCATC
AAAATATTGTTTCAAATGCTGCTGCCTTTCTCAAATGTGTGGAGCATGCTTATGAGCCCACTCCTGATGAT
GTGGCCATATCCTACCTCCCTCTGGCTCATATGTTTGGAGGATTGTACAGGCTGTTGTGTACAGCTGTGG
AGCCAGAGTTGGATTCTTCCAAGGGATATTGCTGGCTGACGACATGAAGACTTTGAAGCCACAT
TGTTTTCCCGCGGTGCCTCGACTCCTTAACAGGATCTACGATAAGGTACAAAATGAGGCCAAGACACCCTTG
AAGAAGTTCTTGTGAAGCTGGCTGTTTCCAGTAAATTCAAAGAGCTTCAAAGGGTATCATCAGGCATGA
TAGTTTCTGGGACAAGCTCATCTTTGCAAAGATCCAGGACAGCCTGGGCGGAAGGGTTCGTGTAATTGTCA
CTGGAGCTGCCCCCATGTCCACTTCAGTCATGACATTCTTCCGGGCAGCAATGGGATGTGAGGTGTATGAA
GCTTATGGTCAAACAGAATGCACAGGTGGCTGTACATTTACATTACCTGGGGACTGGACATCAGGTCACGT
TGGGGTGCCCCCTGGCTTGCAATTACGTGAAGCTGGAAGATGTGGCTGACATGAACACTTTACAGTGAATA
ATGAAGGAGAGGTCTGCATCAAGGGTACAAACGTGTTCAAAGGATACCTGAAGGACCCTGAGAAGACACAG
GAAGCCCTGGACAGTGTGCTGGCTTACACAGGAGACATTGGTCTGCTGGCTCCCGAATGGAACCTCTGAA
GATCATCGACCGTAAAAAGAACATTTTCAAGCTGGCCCAAGGAGAATACATTGCACCAGAGAAGATAGAAA
ATATCTACAACAGGAGTCAACCAGTGTACAAATTTTGTACACGGGGAGAGCTTACGGTCATCCTTAGTAGGA
GTGGTGGTTCTTGACACAGATGTACTTCCCTCATTTGCAGCCAAGCTTGGGGTGAAGGGCTCCTTTGAGGA
ACTGTGCCAAAACCAAGTTGTAAGGGAAGCCATTTTAGAAGACTTGCAAAAATTTGGGAAAGAAAGTGGCC
TTAAAACTTTTGAACAGGTCAAAGCCATTTTCTTCATCCAGAGCCATTTTCCATTGAAAATGGGCTCTTG
ACACCAACATTGAAAGCAAAGCGAGGAGAGCTTTCCAATACTTTTCGACCCAAATTGACAGCCTGTATGA
GCACATCCAGGATTAGGATAAGGTACTTAAGTACCTGCCGGCCCACTGTGCACTGCTTGTGAGAAAATGGA
TTAAAACTATTTCTTACATTTGTTTTGCCTTTCTCCTATTTTTTTTAACTGTAACTCTAAAGCCAT
AGCTTTTGTTTTATATTGAGACATATAATGTGTAACTTAGTTCCCAATAAATCAATCCTGTCTTTCCCA
TCTTCGATGTTGCTAATATTAAGGCTTCAGGGCTACTTTTATCAACATGCCTGTCTTCAAGATCCCAGTTT
ATGTTCTGTGTCCTTCTCATGATTTCCAACCTTAATACTATTAGTAACCACAAGTTCAAGGGTCAAAGGG
ACCCTCTGTGCCTTCTTCTTTGTTTTGTGATAAACATAAAGTGGCCAACAGTCTCTATGCTTATTTACATCT
TCTACTGTTCAAACCTAAGAGATTTTAAATTTCTGAAAACTGCTTACAATTCATGTTTTCTAGCCACTCCA
CAAACCACTAAAATTTTAGTTTTAGCCTATCACTCATGTCAATCATATCTATGAGACAAATGTCTCCGATG
CTCTTCTGCGTAAATTAATTTGTGTACTGAAGGGAAAAGTTGATCATACCAAACATTTTCTTAACTCTCT
AGTTAGATATCTGACTTGGGAGTATTAATAATTTGGGTCTATGACATACTGTCCAAAAGGAATGCTGTCTT
AAAGCATTATTTACAGTAGGAAGTGGGGAGTAAATCTGTTCCCTACAGTTTGTCTGTGAGCTGGAAGCTGT
GGGGGAAGGAGTTGACAGGTGGGCCCACTGAACTTTTCCAGTAAATGAAGCAAGCACTGAATAAAAACTC
CTGAAGTGGGAACAAAGATCTACAGGCAAGCAAGATGCCACACAACAGGCTTATTTTCTGTGAAGGAACC
AACTGATCTCCCCACCCTTGATTAGAGTTCTGTCTACCTTACCCACAGATAACACATGTTGTTTCTA
CTTGTAATGTAAAGTCTTTAAATAAACTATTACAGATAAAAAA

4/6

FIGURE 4

MDALKPPCLWRNHERGKKDRDSCGRKNSEPGSPHSLEALRDAAPSQGLNFFLLFTKMLFIFNFLFSPLPTP
ALICILTFGAIFLWLI TRPQPVLP LLDLNNQSVGIEGGARKGVSQKNNDLTSCCFSDAKTMYEVFQRGLA
VSDNGPCLGYRKPNQPYRWLSYKQVSDRAEYLGSCLLHKGYKSSPDQFVGIFAQNRPEWIISELACYTYSM
VAVPLYDTLGPEAIVHIVNKADIAMVICDTPQKALVLIGNVEKGFTPSLKVIIIMDPFDDDLKQERGEKSGI
EILSLYDAENLGKEHFRKPVPPSPEDLSVICFTSGTTGDPKGAMITHQNIIVSNAAAFKCEVHAYEPTPDD
VAISYLP LAHMFERIVQAVVYSCGARVGFFQGDIRLLADDMKTLKPTLFPAVPRLNRIYDKVQNEAKTPL
KKFLLKLAVSSKFELQKGIIRHDSFWDKLI FAKIQDSLGGVRVIVTGAAPMSTSVMTFFRAAMGCQVYE
AYGQTECTGGCTFTLP GDWTS GHVGVPLACNYVKLEDVADMN YFTVNNEGEVCIKGTNVFKGYLKDPEKTQ
EALDS DGWLHTGDI GRWLPNGTLKI IDRKNIFKLAQGEYIAPEK IENIYNRSQPVLQIFVHGESLRSSLV
GVVVPD TDVLP SF AAKLG VKGSFEELCQNQVVREAILEDLQKIGKESGLKTFEQVKAIFLHPEPF SIENGL
LTPTLKAKRGELSKYFRTQIDSLYEHIQD

Type II transmembrane domain:

Amino acids 61-80

Putative AMP-binding domain signature:

Amino acids 314-326

N-glycosylation sites:

Amino acids 102-106;588-594;
619-623

5/6

FIGURE 5

CAACCATGCAAGGACAGGGCAGGAGAAGAGGAACCTGCAAAGACATATTTTGTTCCTGCAAAAATGGCATCTTAC
CTTTATGGAGTACTCTTTGCTGTTGGCCTCTGTGCTCCAATCTACTGTGTGTCCCCGGCCAATGCCCCCAG
TGCATACCCCCGCCCTTCCCTCCACAAAGAGCACCCCTGCCTCACAGGTGTATTCCCTCAACACCCGACTTTG
CCTTCCGCCTATACCGCAGGCTGGTTTTGGAGACCCCGAGTCAGAACATCTTCTTCTCCCTGTGAGTGTG
TCCACTTCCCTGGCCATGCTCTCCCTTGGGGCCCACTCAGTCACCAAGACCCAGATTCTCCAGGGCCTGGG
CTTCAACCTCACACACACACCAGAGTCTGCCATCCACCAGGGCTTCCAGCACCTGGTTCACTCACTGACTG
TTCCAGCAAAGACCTGACCTTGAAGATGGGAAGTGCCCTCTTCGTCAAGAAGGAGCTGCAGCTGCAGGCA
AATTCTTGGGCAATGTCAAGAGGCTGTATGAAGCAGAAGTCTTTTCTACAGATTTCTCCAACCCCTCCAT
TGCCCAGGCGAGGATCAACAGCCATGTGAAAAAGAAGACCCAAGGGAAGGTTGTAGACATAATCCAAGGCC
TTGACCTTCTGACGGCCATGGTTCTGGTGAATCACATTTTCTTTAAAGCCAAGTGGGAGAAGCCCTTTCAC
CTTGAATATACAAGAAAGAACTTCCCATTCTGCTGGTGGGCGAGCAGGTCACTGTGCAAGTCCCCATGATGCA
CCAGAAAGAGCAGTTCGCTTTTGGGGTGGATACAGAGCTGAACTGCTTTGTGCTGCAGATGGATTACAAGG
GAGATGCCGTGGCCTTCTTTGTCTCCCTAGCAAGGGCAAGATGAGGCAACTGGAACAGGCCTTGTGAGCC
AGAACACTGATAAAGTGGAGCCACTCACTCCAGAAAAGGTGGATAGAGGTGTTTATCCCCAGATTTTCCAT
TTCTGCCTCCTACAATCTGGAAACCATCTCCCGAAGATGGGCATCCAAAATGCCTTTGACAAAAATGCTG
ATTTTTCTGGAATTGCAAAGAGAGACTCCCTGCAGGTTTCTAAAGCAACCCACAAGGCTGTGCTGGATGTGAGT
GAAGAGGGCACTGAGGCCACAGCAGCTACCACCACCAAGTTCATAGTCCGATCGAAGGATGGTCCCTCTTA
CTTCACTGTCTCCTTCAATAGGACCTTCCCTGATGATGATTACAAATAAAGCCACAGACGGTATTCTCTTTC
TAGGGAAAGTGGAAAATCCCACTAAATCCTAGGTGGGAAATGGCCTGTAACTGATGGCACATTGCTAATG
CACAAGAAATAACAAACCACATCCCTCTTTCTGTTCTGAGGGTGCATTTGACCCAGTGGAGCTGGATTGCG
CTGGCAGGGATGCCACTTCCAAGGCTCAATACCAAACCATCAACAGGGACCCAGTCACAAGCCAACACC
CATTACCCCAAGTCAGTGCCCTTTTCCACAAATTCTCCAGGTAAGTCTGCTTCAATGGGATGTTGCTGGGTT
ACCATATTTCCATTCTTGGGGCTCCAGGAATGGAAATACGCCAACCCAGGTAGGCACCTCTATTGCAG
AATTACAATAACACATTCAATAAACTAAATATGAATTCAAAAA
AA

6/6

FIGURE 6

MASYLYGVLFVAVGLCAPIYCVSPANAPSAYPRPSSTKSTPASQVYSLNTDFAFRLYRRLVLETSPSONIFFS
PVSVSTSLAMLSLGAHSVTKTQILQGLGFNLTHTPESAIHQGFQHLVHSLTVPSKDLTLKMGSALFVKKEL
QLQANFLGNVKRLYEAEVFSTDFSNPSIAQARINSHVKKKTQGKVVDIIQGLDLLTAMVLVNHIFFKAKWE
KPFHLEYTRKNFPFLVGEQVTVQVPMMHQKEQFAFGVDTELNCFVLQMDYKGDVAFFVLPSKGKMRQLEQ
ALSARTLIKWSHSLQKRWIEVFIPRFSISASYNLETILPKMGIONAFDKNADFSGIAKRDSLQVSKATHKA
VLDVSEEGTEATAATTTKFIVRSKDGPSYFTVSNRTFLMMITNKATDGILFLGKVENPTKS

Signal peptide:	Amino acids 1-20
N-glycosylation sites:	Amino acids 101-105;390-394
Tyrosine kinase phosphorylation site:	Amino acids 377-385
N-myristoylation sites:	Amino acids 7-13;97-103;326-332; 363-369

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
12 July 2001 (12.07.2001)

PCT

(10) International Publication Number
WO 01/49715 A3

- (51) International Patent Classification⁷: **C12N 15/12**,
C07K 14/71, A61K 38/18, 39/395, C12N 15/62, C07K
16/28
- (71) Applicant (for all designated States except US): **GENEN-
TECH, INC.** [US/US]; 1 DNA Way, South San Francisco,
CA 94080-4990 (US).
- (21) International Application Number: PCT/US00/30952
- (72) Inventors; and
(75) Inventors/Applicants (for US only): **ASHKENAZI,**
Avi, J. [US/US]; 1456 Tarrytown Street, San Mateo,
CA 94402 (US). **GODDARD, Audrey** [CA/US]; 110
Congo Street, San Francisco, CA 94131 (US). **GURNEY,**
Austin, L. [US/US]; 1 Debbie Lane, Belmont, CA 94002
(US). **NAPIER, Mary, A.** [US/US]; 1015 Hayne Road,
Hillsborough, CA 94010 (US). **WATANABE, Colin, K.**
[US/US]; 128 Corliss Drive, Moraga, CA 94556 (US).
WOOD, William, I. [US/US]; 35 Southdown Court,
Hillsborough, CA 94010 (US).
- (22) International Filing Date:
8 November 2000 (08.11.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
- | | | |
|----------------|-------------------------------|----|
| PCT/US00/00376 | 6 January 2000 (06.01.2000) | US |
| PCT/US00/04342 | 18 February 2000 (18.02.2000) | US |
| PCT/US00/05841 | 2 March 2000 (02.03.2000) | US |
| PCT/US00/08439 | 30 March 2000 (30.03.2000) | US |
| PCT/US00/20710 | 28 July 2000 (28.07.2000) | US |
- (74) Agents: **BARNES, Elizabeth M.** et al.; c/o Genentech,
Inc., MS49, 1 DNA Way, South San Francisco, CA 94080-
4990 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,

[Continued on next page]

(54) Title: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH

MTSPSLLLLLLPPLLLGAFPPAAAGPFPKMDKVVPQVARLGRTVRLQCPVEGDPPLTMWTKDGRTH
SGWSRFRVLPGQLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVLDLDDISPGKESLGPDSSSGGQEDPASQ
QWARPRFTQPSKMRRRVIAIPVGSVRLKCVASGHPRPDITWMKDDQALTRPEAAEPRKKKWTLSLKNLRP
EDSGKYTCRVSNRAGAINATYKVDVIQRTSRKPVLTGTHPVNTTVDFGGTTSFOCKVRSVDKPFVIQWLKRV
EYGAERHNSITIDVGGQKFVVLPTGDIVSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRSAFLT
LPDPKPPGPPVASSSSATSLPWPVIGIPAGAVFILGTLTLLWLCQAQKKPCTPAPAPPLPGHRPFGTARDR
SGDKDLPSLAALSAGPGVGLCEEHGSAPAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTSHSHVEGKVV
QHIHYQC

Signal peptide: Amino acids 1-17

Putative transmembrane domain: Amino acids 376-396

Tyrosine kinase phosphorylation sites: Amino acids 212-220;329-337

N-glycosylation sites: Amino acids 111-115;231-235;255-259;
293-297

Immunoglobulins and major histocompatibility protein:
Amino acids 219-237

cAMP- and cGMP-dependent protein kinase phosphorylation site:
Amino acids 202-206

N-myristoylation sites: Amino acids 95-101;228-234;261-267;
317-323;334-340;382-388;443-449

(57) Abstract: The present invention concerns methods and compositions for inhibiting neoplastic cell growth. In particular, the present invention concerns antitumor compositions and methods for the treatment of tumors. The invention further concerns screening methods for identifying growth inhibitory, e.g., antitumor compounds. The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

WO 01/49715 A3



DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

- (84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

- (88) **Date of publication of the international search report:**
4 April 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/30952

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/71 A61K38/18 A61K39/395 C12N15/62
C07K16/28

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 63088 A (BAKER KEVIN ;CHEN JIAN (US); GENENTECH INC (US); YUAN JEAN (US); G) 9 December 1999 (1999-12-09) * see particularly the passages relating to PR0943 *	1-7,13, 15-17, 19-40
A	WO 91 15510 A (SQUIBB BRISTOL MYERS CO) 17 October 1991 (1991-10-17) the whole document	
A	EP 0 315 289 A (ONCOGEN) 10 May 1989 (1989-05-10) the whole document	
A	WO 95 09005 A (SQUIBB BRISTOL MYERS CO) 6 April 1995 (1995-04-06) the whole document	
	--- -/--	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Z" document member of the same patent family

Date of the actual completion of the international search

2 August 2001

Date of mailing of the international search report

26. 10. 01

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Smalt, R

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/30952

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X, L	WO 00 53755 A (BAKER KEVIN P ; GENENTECH INC (US); ASHKENAZI AVI J (US); GODDARD A) 14 September 2000 (2000-09-14) * see particularly the passages relating to PR0943 * L: priority ---	1-40
E,L	WO 00 73454 A (GENENTECH INC) 7 December 2000 (2000-12-07) * see particularly the passages relating to PR0943 * L: priority ---	1-7,13, 15-17, 19-40
E,L	WO 01 00661 A (GENSET SA ; BEJANIN STEPHANE (FR); BIHAIN BERNARD (US); BOUR BARBAR) 4 January 2001 (2001-01-04) * see particularly the passages relating to PR0943 * L: priority ---	1-7,13, 15-17, 19-40
E,L	WO 01 09327 A (BAKER KEVIN P ; GENENTECH INC (US); ASHKENAZI AVI J (US); GODDARD A) 8 February 2001 (2001-02-08) * see particularly the passages relating to PR0943 * L: priority -----	1-7,13, 15-17, 19-40

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/30952

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claim 12, and claims 8-10 in as far as they pertain to in vivo use, are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☒ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Claims 1-40, all partially.

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Present claims 1-8, 10-14, and 16-18 relate to a compound defined by reference to a desirable characteristic or property, namely acting as an agonists towards the polypeptides of claim 1.

The claims cover all products having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such products. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product/compound/method/apparatus by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to agonistic antibodies directed at the polypeptides of claim 1.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-40, all partially

PR0943 protein and nucleic acid encoding it, as represented by seq.ID's 2 and 1, respectively, vector comprising the nucleic acid, host containing the vector, and process for producing the protein using said hosts. Chimeric molecules comprising said protein, antibody against said protein, composition comprising said protein or said antibody for use against tumour cell growth.

2. Claims: 1-40, all partially

PR01250 protein and nucleic acid encoding it, as represented by seq.ID's 4 and 3, respectively, vector comprising the nucleic acid, host containing the vector, and process for producing the protein using said hosts. Chimeric molecules comprising said protein, antibody against said protein, composition comprising said protein or said antibody for use against tumour cell growth.

3. Claims: 1-40, all partially

PR01337 protein and nucleic acid encoding it, as represented by seq.ID's 6 and 5, respectively, vector comprising the nucleic acid, host containing the vector, and process for producing the protein using said hosts. Chimeric molecules comprising said protein, antibody against said protein, composition comprising said protein or said antibody for use against tumour cell growth.

INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/US 00/30952

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9963088	A	09-12-1999	AU 4328699 A	20-12-1999
			AU 2212299 A	26-07-1999
			WO 9935170 A	15-07-1999

WO 9115510	A	17-10-1991	US 5416192 A	16-05-1995
			AT 162549 T	15-02-1998
			AU 639787 B	05-08-1993
			AU 7744791 A	30-10-1991
			CA 2079218 A	04-10-1991
			CN 1058782 A	19-02-1992
			CS 9100917 A	17-06-1992
			DE 69128779 D	26-02-1998
			EP 0527806 A	24-02-1993
			FI 924444 A	02-10-1992
			GR 3026263 T	29-05-1998
			HU 63438 A	30-08-1993
			IE 911117 A	09-10-1991
			NO 923861 A	01-12-1992
			NZ 237661 A	27-07-1993
			OA 9711 A	30-08-1993
			PT 97303 A	30-09-1992
			US 5885961 A	23-03-1999
			US 5965723 A	12-10-1999
			ZA 9102473 A	30-12-1992

EP 0315289	A	10-05-1989	AU 629185 B	01-10-1992
			AU 2474088 A	11-05-1989
			CN 1040926 A	04-04-1990
			DK 618188 A	07-05-1989
			FI 885055 A	07-05-1989
			IL 88276 A	18-08-1993
			JP 2072869 A	13-03-1990
			KR 9102520 B	23-04-1991
			NO 884943 A	08-05-1989
			NZ 226799 A	27-08-1991
			PT 88943 A	30-11-1989
			ZA 8808269 A	26-07-1989

WO 9509005	A	06-04-1995	US 5814307 A	29-09-1998
			AU 7844794 A	18-04-1995
			CA 2172808 A	06-04-1995
			EP 0721344 A	17-07-1996
			JP 9505560 T	03-06-1997

WO 0053755	A	14-09-2000	AU 1748200 A	19-06-2000
			AU 1749900 A	12-07-2000
			AU 2192800 A	12-07-2000
			AU 2224800 A	28-09-2000
			AU 2399300 A	28-09-2000
			AU 2495200 A	28-09-2000
			AU 2596700 A	28-09-2000
			AU 2600800 A	28-09-2000
			AU 2883600 A	28-09-2000
			AU 3072199 A	27-09-1999
			AU 3107700 A	28-09-2000
			AU 3381600 A	28-09-2000
			AU 3514400 A	28-09-2000
			EP 1135485 A	26-09-2001

INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/US 00/30952

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0053755 A		EP 1141284 A	10-10-2001
		EP 1141289 A	10-10-2001
		WO 0053753 A	14-09-2000
		WO 0053754 A	14-09-2000
		WO 0153486 A	26-07-2001
		WO 0053756 A	14-09-2000
		WO 0053757 A	14-09-2000
		WO 0053758 A	14-09-2000
		WO 0032221 A	08-06-2000
		WO 0053750 A	14-09-2000
		WO 0037638 A	29-06-2000
		WO 0037640 A	29-06-2000
		WO 0053751 A	14-09-2000
		WO 0053752 A	14-09-2000
		AU 1932000 A	03-07-2000
		AU 2215300 A	28-12-2000
		AU 3246100 A	28-12-2000
		AU 3774300 A	18-12-2000
		AU 4328699 A	20-12-1999
		AU 5441200 A	18-12-2000
		AU 5460100 A	18-12-2000
		AU 5591100 A	18-12-2000
		EP 1141285 A	10-10-2001
		WO 0075327 A	14-12-2000
		WO 0073454 A	07-12-2000
		WO 0073445 A	07-12-2000
		WO 0073348 A	07-12-2000
		WO 0073452 A	07-12-2000
		WO 0036102 A	22-06-2000
		WO 0075316 A	14-12-2000
		AU 2883700 A	09-01-2001
		WO 0078961 A	28-12-2000
		AU 2883900 A	30-01-2001
		WO 0104311 A	18-01-2001
		AU 1749800 A	04-10-2000
WO 0073454 A	07-12-2000	AU 1748200 A	19-06-2000
		AU 1932000 A	03-07-2000
		AU 2192800 A	12-07-2000
		AU 2215300 A	28-12-2000
		AU 2495200 A	28-09-2000
		AU 2600800 A	28-09-2000
		AU 3246100 A	28-12-2000
		AU 3381600 A	28-09-2000
		AU 3514400 A	28-09-2000
		AU 3774300 A	18-12-2000
		AU 4328699 A	20-12-1999
		AU 5441200 A	18-12-2000
		AU 5460100 A	18-12-2000
		AU 5591100 A	18-12-2000
		EP 1135485 A	26-09-2001
		EP 1141285 A	10-10-2001
		EP 1141289 A	10-10-2001
		WO 0053753 A	14-09-2000
		WO 0053755 A	14-09-2000
		WO 0153486 A	26-07-2001
		WO 0075327 A	14-12-2000
		WO 0053757 A	14-09-2000

INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/US 00/30952

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0073454 A		WO 0053758 A	14-09-2000
		WO 0073445 A	07-12-2000
		WO 0073348 A	07-12-2000
		WO 0073452 A	07-12-2000
		WO 0032221 A	08-06-2000
		WO 0036102 A	22-06-2000
		WO 0037640 A	29-06-2000
		WO 0075316 A	14-12-2000
		AU 2883600 A	28-09-2000
		AU 2883700 A	09-01-2001
		WO 0053756 A	14-09-2000
		WO 0078961 A	28-12-2000
		AU 2883900 A	30-01-2001
		WO 0104311 A	18-01-2001
		AU 1749800 A	04-10-2000
		AU 1749900 A	12-07-2000
		AU 2390700 A	05-02-2001
		AU 2399300 A	28-09-2000
		AU 3107000 A	19-06-2000
		AU 4011300 A	05-02-2001
		EP 1141284 A	10-10-2001
		EP 1135491 A	26-09-2001
		WO 0105972 A	25-01-2001
		WO 0032778 A	08-06-2000
		WO 0055319 A	21-09-2000
		WO 0037638 A	29-06-2000
		WO 0105836 A	25-01-2001
WO 0100661 A	04-01-2001	AU 1675900 A	12-07-2000
		AU 5559400 A	31-01-2001
		AU 5838200 A	31-01-2001
		AU 5838300 A	31-01-2001
		EP 1144444 A	17-10-2001
		WO 0100806 A	04-01-2001
		WO 0100803 A	04-01-2001
		WO 0037491 A	29-06-2000
WO 0109327 A	08-02-2001	AU 2883900 A	30-01-2001
		AU 3514400 A	28-09-2000
		AU 3774300 A	18-12-2000
		AU 5441200 A	18-12-2000
		AU 5460100 A	18-12-2000
		AU 6391000 A	19-02-2001
		WO 0153486 A	26-07-2001
		WO 0104311 A	18-01-2001
		WO 0053758 A	14-09-2000
		WO 0073454 A	07-12-2000
		WO 0073445 A	07-12-2000
		WO 0073452 A	07-12-2000
		AU 1747100 A	17-04-2001
		AU 1748200 A	19-06-2000
		AU 2390700 A	05-02-2001
		AU 5920099 A	03-04-2000
		EP 1112361 A	04-07-2001
		EP 1135485 A	26-09-2001
		WO 0119987 A	22-03-2001
		WO 0032221 A	08-06-2000
		WO 0105836 A	25-01-2001

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/30952

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0109327 A		AU 1749800 A	04-10-2000
		AU 1749900 A	12-07-2000
		AU 2192800 A	12-07-2000
		AU 2215300 A	28-12-2000
		AU 2600800 A	28-09-2000
		AU 3381600 A	28-09-2000
		AU 3878400 A	10-04-2001
		AU 5152700 A	02-01-2001
		AU 5922999 A	03-04-2000
		AU 7573000 A	26-03-2001
		EP 1114152 A	11-07-2001
		EP 1141284 A	10-10-2001
		EP 1141289 A	10-10-2001
		WO 0053753 A	14-09-2000
		WO 0053757 A	14-09-2000
		WO 0118210 A	15-03-2001
		WO 0077037 A	21-12-2000
		WO 0116318 A	08-03-2001
		WO 0055319 A	21-09-2000
		WO 0037638 A	29-06-2000
		WO 0037640 A	29-06-2000
		WO 0075316 A	14-12-2000
		AU 2224800 A	28-09-2000
		AU 2495200 A	28-09-2000
		AU 2596700 A	28-09-2000
		AU 2883600 A	28-09-2000
		AU 2883700 A	09-01-2001
		AU 3107700 A	28-09-2000

24

AR

Identification of centerin: a novel human germinal center B cell-restricted serpin

XP-002229972

J. Kimble Frazer³, Deborah G. Jackson^{1,2}, Jean-Phillipe Gaillard², Michael Lutter¹, Yong-Jun Liu⁴, Jacques Banchereau², J. Donald Capra³ and Virginia Pascual^{1,2}

¹ UT Southwestern Medical Center at Dallas, Dallas, USA

² Baylor Institute for Immunology Research, Dallas, USA

³ Oklahoma Medical Research Foundation, Oklahoma City, USA

⁴ DNAX Institute for Molecular and Cellular Biology, Palo Alto, USA

P.D. 00-10-2000	10
P. 3039-3048	

For naive B cells to mature in response to antigen triggering and become either plasma cells or memory B cells, a complex array of events takes place within germinal centers (GC) of secondary lymphoid organs. With the long-term objective of defining and characterizing molecules that control the generation of GC, we have subtracted RNA messages derived from highly purified B cells at the follicular mantle stage of differentiation from GC B cells. Using this approach, we have identified a novel molecule, centerin, belonging to the family of serine-protease inhibitors or serpins. Transcription of centerin is highly restricted to GC B cells and their malignant counterparts, Burkitt's lymphoma lines. The putative centerin protein shares the highest sequence identity with thyroxine-binding globulin and possesses arginine/serine at its P1/P1' active site, suggesting that it interacts with a trypsin-like protease(s). In addition, several other sequence features of centerin also indicate that it serves as a bonafide protease inhibitor. Finally, we demonstrate differentially up-regulated transcription of this novel gene by resting, naive B cells stimulated *in vitro* via CD40 signaling, while *Staphylococcus aureus* Cowan strain-mediated B cell activation fails to generate this response. Because CD40 signaling is required for naive B cells to enter the GC reaction and for GC B cells to survive, it is likely that centerin plays a role in the development and/or sustaining of GC.

Received	10/1/00
Revised	15/5/00
Accepted	28/6/00

Key words: Serpin / Germinal center / B lymphocyte / CD40 ligand activation

1 Introduction

After leaving the bone marrow, B cells are uniquely able to modify their Ig variable region genes through the processes of isotype switching, somatic mutation, and receptor editing, all of which contribute to the generation of high-affinity antibodies. It is widely accepted that these genetic alterations occur as B cells passage through the germinal center (GC), a microanatomical site located within the secondary follicles of peripheral lymphoid organs. B cells entering the GC reaction undergo a series of phenotypic and functional changes that ultimately lead to the generation of terminally differentiated antibody-secreting plasma cells and memory B cells [1–5].

[1 20448]

J. K. Frazer and D. G. Jackson contributed equally to this work.

Abbreviations: GC: Germinal center FM: Follicular mantle
RDA: Representational difference analysis **TBG:** Thyroxine-binding globulin **RPA:** RNase protection assay **PCI:** Protein C inhibitor

GC B cells can be purified from human tonsils using antibodies against cell surface markers and flow cytometry. Staining with anti-IgD and anti-CD38 antibodies allows the separation of four mutually exclusive tonsillar B cell populations; (1) single-positive IgD cells corresponding to follicular mantle (FM) cells or Bm1+Bm2; (2) single-positive CD38 cells corresponding to germinal center (GC) cells or Bm3+Bm4; (3) double-negative B cells corresponding to the memory population (Bm5), and (4) double-positive cells representing a combination of cells, including those at a transitional stage between FM and GC (Bm2'), and IgD + GC cells (reviewed in [4, 5]). The availability of these pure B cell populations provides a unique substrate for isolating genes restricted to the GC. We used representational difference analysis (RDA) to subtract FM from GC RNA [6, 7] and obtained a series of novel RNA messages, one of which represents a new member of the serine protease inhibitor (serpin) superfamily [8–10].

Serpin act as physiological inhibitors of specific proteases with which they form stable inactive complexes.

© WILEY-VCH Verlag GmbH, D-69451 Weinheim, 2000

0014-2980/00/1010-3039\$17.50 + .50/0

Although serpins bind to their protease target with high affinity, most of them have the ability to inhibit more than one protease (cross-class inhibition). In general, serpins contain a highly exposed reactive center loop near the C terminus of the molecule that mimics an ideal protease substrate. The inhibitory specificity of the serpin is in part defined by one amino acid residue (P1) in the reactive site loop. For example, proteases of trypsin-like specificity attack peptide bonds immediately after the P1 residues arginine and lysine, whereas chymotrypsin cleaves substrates with large hydrophobic amino acids in the P1 position. Interaction of the serpin with the active site of its target protease triggers conformational changes and results in an irreversible serpin-protease complex that can be detected by SDS-polyacrylamide gel electrophoresis. The hinge region located at a position N-terminal (P_{17} - P_8) to the cleavage site (P_1 - P_1') of serpin proteins is believed to permit the mobility necessary to mediate inhibition of serine proteases [8–16]. In this way, serpins play crucial roles in regulating proteolytic enzymes that are involved in a wide variety of vital processes including blood coagulation, fibrinolysis, complement activation, inflammation, cell migration, and apoptosis.

Transcription of the novel serpin molecule that we describe here is highly restricted to GC B cells. Its translated amino acid product shares the highest identity with thyroxine-binding globulin (TBG) [17, 18] and displays arginine/serine at its P1/P1' active site, suggesting that it interacts with trypsin-like protease(s) to mediate its biological function(s). We show that it is possible to up-regulate the transcription of this gene in resting, naive B cells through CD40 signaling, while *Staphylococcus aureus* Cowan strain (SAC) activation fails to do so. The implications of these findings are discussed below.

2 Results

2.1 Representational difference analysis of FM and GC human B cell populations

Using cDNA RDA [19], FM and GC message pools were subtracted from one another to identify specific mRNA transcriptionally up- or down-regulated in the FM-to-GC transition. Because the double-sorted cell populations were available in very limited quantities (*i.e.* 2×10^5 FM or GC cells), variations to standard cDNA RDA protocols were engineered into the procedure, as described elsewhere [7]. After three rounds of subtractive hybridization and differential amplification with progressively higher subtraction ratios, a large number of distinct sequence fragments were obtained in both the FM-GC (FM-specific genes) and GC-FM (GC-specific genes) subtractions. These were cloned, sequenced, and searched

against NCBI-administered NR and EST databases using the BLAST algorithm [20]. In the GC-specific direction, 91 unique gene fragments were isolated. A 245-bp fragment corresponding to the CD38 cDNA was obtained and represented 4% of the total number of independent isolates. CD38 was one of the two surface markers used to classify B cells by flow cytometry as being of either FM or GC origins. Therefore, this finding provided confirmation of the successful execution of RDA.

2.2 Isolation and cloning of centerin

One RDA fragment representing 12% of the total number of sequenced fragments, was a 612-bp sequence with no identity at the nucleotide level to any entries in either NR or EST databases. Putative amino acid sequence alignments, however, disclosed significant identity between the translated amino acid sequence of the RDA fragment and several members of the serpin family. cDNA extension by tailing and/or tagging the first cDNA strand followed by amplification of the ends with at least one internal primer (RACE) permitted us to extend the initial sequence in both directions from the 612 to 1,500 bp. We named this gene "centerin". Upon cloning of the 5' end, two distinct centerin messages corresponding to alternatively spliced isoforms were identified. The longest clone contains two putative starting sites (translation starting at M1 or M2), one of which (M2) is followed by a putative 5' leader sequence. The second form (translation starting at M3) displays a 190-bp deletion that removes the leader and the contiguous N terminus (Fig. 1). The deleted region encodes the first α -helix of the serpin molecule according to the available three-dimensional structure of other serpin molecules [11, 21–24]. Because this first α -helix is essential for the serpin molecule to fold appropriately prior to and after interaction with its target protease, it is possible that transcripts starting at M3 may not be functional (Fig. 1).

Overall, the two cloned centerin messages encode open reading frames (ORF) of 417 and 337 amino acids respectively. As Fig. 1 shows, the presence of an additional methionine residue (M1) 18 amino acids upstream of, and in frame with, the leader's first methionine (M2) suggests the possibility of a larger, 435-residue protein being made as well.

2.3 Centerin encodes a potentially active serpin

Centerin displays the closest identity (50% at the amino acid level) with the non-inhibitory serpin TBG precursor [17, 18]. However, the percentage of identity between

Leader peptide
 ←-----→

MRGQRRRGTCCKDIFCSK**M**ASYLYGVLFVAVGLCAPIYCVSPANAPSAYPRPSSIKSTPAS 60
 VYSLNTDFAFRLYRRLVLETPSQNIFFSPVSVSTSLAML~~SLGAHSVTKTQILQGLGFNLT~~ 120
 HTPESAIOHGFQHLVHSLTVPSKDLTLKMGSAFLVKKELQLQANFLGNVKRLYEAEVFST 180
 DFSNPISIAQARINSHVKKKTQGVVDIIQGLDLTAMVLVNHFFKAKWEKPFHLEYTRKN 240
 FPFLVGEQVTVQVPMHKEQFAFGVDTELNCFVLQMDYKGDVAFFVLPSKGMKMRQLEQ 300
 ALSARTLIKWSHSLQKRWIEVFIPRFSISASYNLETILPKMGIQNNAFDKNADFSGI~~AKR~~ 360
 DSLQVSKATHKAVLDVSEEGTEATAATTTK**FIV****R**SKDGPSYFT**V****SFNR****TFL****MM**ITNKATD 420
 GILFLGKVENPTKS 434

P1 P1'

Fig. 1. Translated amino acid sequence of the centerin gene. Bold methionines (**M**) represent potential translation initiation sites. The underlined sequence at the N terminus is deleted from the second centerin splice variant. This deletion removes the second M start site and creates a frame shift such that the first M is no longer in frame with the rest of the sequence. A putative 19-amino acid leader peptide sequence is represented with a double-headed arrow. The P1 and P1' residues within the serpin active site are bold and double underlined. The serpin signature sequence at the C terminus is bold and underlined. The underlined valine (V) residue at position 155 is deleted in one allele of the Burkitt's cell line Raji.

these two molecules around the hinge region and active site is only 20 % (Fig. 2). At its C terminus, centerin contains all the motifs necessary to be considered an active serpin, i.e. (a) "serpin signature sequence", (b) hinge region, and (c) active site composed of P1 (arginine) and P1' (serine) residues (Figs. 1 and 2) [15].

2.4 Centerin transcription is highly to germinal center B cells

To test its differential expression, we performed Northern blot analyses, RNase protection assays (RPA), PCR amplification, and *in situ* hybridization experiments using primers/probes common to the two possible centerin isoforms. The results of these experiments are depicted in Figs. 3–6. Northern blotting with the RDA-isolated centerin gene fragment demonstrated that this serpin

was largely restricted in its transcription (Fig. 3). A survey of eight normal human tissues (heart, brain, placenta, lung, liver skeletal muscle, kidney, and pancreas) failed to detect any signal in repetitive experiments with appropriate positive controls (data not shown). Assaying eight human cancer cell lines, only the Burkitt's lymphoma Raji and the HL-60 promyelocytic lines were positive for the 2.0-kb message (Fig. 3). Burkitt's lymphoma corresponds to a transformed centroblast phenotype in terms of differentiation stage. Thus, the presence of the centerin message in this line is consistent with its temporal regulation within the B cell lineage. An additional Northern blot examining RNA from six different human immunological tissues revealed the presence of the centerin

↓
P1 P1'

Centerin	LDVSEEGTEATAATTTK FIV	R SKDGPSYFT V SFNR TFL MM
PCI	VE-D-S--R-A---G-I-TF-	-ARLN-QRL- ----P----F
ATIII	-E-N---S--A-S-AVVIAG-	-LN.-NRV-FKA--P--VF
TBG	-HIG-K----A-VPEVELSDQ	PENTFLHPITQID-S-MLL

Fig. 2. Translated amino acid sequence comparison of the active sites of centerin, PCI, anti-thrombin III (ATIII), and TBG. Dashes represent amino acid identities. Periods represent gaps introduced to maximize alignment. P1 and P1' residues are bold.

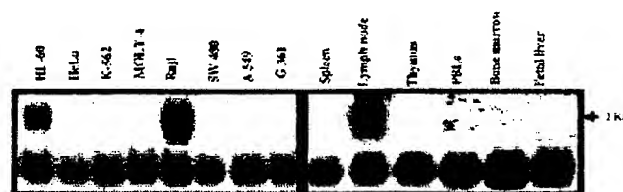


Fig. 3. Northern Blot analysis of cancer cell lines (left) and immune system tissues (right) using a probe corresponding to the 612-bp centerin RDA fragment. Control hybridization with an actin probe is shown in the lower panels. HL-60: promyelocytic leukemia; HeLa: cervical carcinoma; K-562: chronic myelogenous leukemia; MOLT-4: lymphoblastic leukemia; Raji: Burkitt's lymphoma; SW480: colorectal adenocarcinoma; A549: lung carcinoma; G361: melanoma cell line.

message only in the lymph node sample, a site rich in GC B cells (Fig. 3). Even though the centerin message is present in the pro-myelocytic leukemia line HLA-60, the bone marrow and peripheral blood RNA samples failed to yield any signal using the centerin probe in Northern blot analysis, suggesting that myeloid precursors as well as mature monocytes and granulocytes do not transcribe this gene. The centerin message was also negative on mature macrophages by PCR assay (Fig. 4). Although not included in Fig. 4, PCR amplification of unfractionated peripheral blood after red blood cell lysis (70 % granulocytes), B cell-depleted PBMC (T cells, NK cells and monocytes) and myeloid-derived dendritic cell precursors (CD14⁺, CD11c⁺) also yielded negative results. We believe these data support the lack expression of the centerin gene in cells of myeloid lineage.

To confirm the restricted expression of centerin within mature B cells, CD38⁺ blood B cells were purified using antibody-coated magnetic beads. These cells are a mixture of 60 % naive and 40 % memory B cells according to CD27 staining (data not shown). PCR amplification of the cDNA of these cells failed to reveal centerin's transcription, as opposed to CD38⁺ tonsillar GC cells included in the same experiment (Fig. 4). Because plasma cells can only be isolated in low amounts from tonsils, we chose two human myeloma cell line (U266 and RPMI 8221) to assay for centerin transcription. The result of one of these experiments is included in Fig. 4.

Two RNA fragments of 260 and 300 bp, respectively, were protected when we used an antisense centerin probe in RPA (Fig. 5). This is due to the presence of each of the two centerin messages, one of which possesses an internal nucleotide deletion that correlates with a protein product truncated by 97 amino acids (underlined in

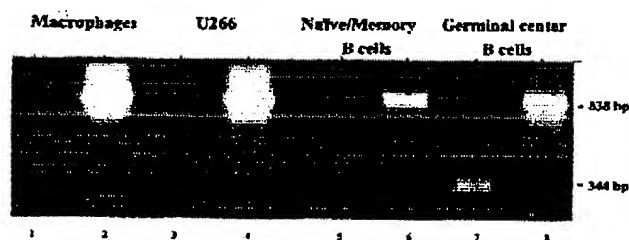


Fig. 4. PCR amplification of the centerin and actin transcripts from human macrophages (lanes 1 and 2), U266 line (human myeloma cell line) (lanes 3 and 4), peripheral blood CD20⁺CD38⁺ naive and memory B cells (lanes 5 and 6), tonsillar CD20⁺CD38⁺ GC B cells (lanes 7 and 8). Bands at 838 bp correspond to the actin product and are visualized in all corresponding lanes, while the 344 bp band corresponding to the centerin product is present only in the GC B cell lane.

Fig. 1), thereby resulting in the smaller protected band. The largest isoform appears to be preferentially transcribed in all positive samples tested so far (Fig. 5 and data not shown). In the case of the Burkitt's line Raji, two additional bands were seen by RPA at 274 and 235 bp, respectively (Fig. 5). Upon cloning and sequencing the centerin message from this cell line, we were able to identify four different transcripts: two identical to those isolated from GC B cells, two displaying a triplet deletion resulting in the loss of an amino acid (valine) at position 155. This finding most likely reflects a deletion in one copy of the centerin gene. As a result, transcripts from Raji would theoretically encode centerin proteins of 417, 416, 337, and 336 amino acids (Fig. 5). While Northern blots cannot distinguish this subtle 3-bp difference (in fact, Northern blots fail to resolve the two centerin isoforms which differ by 190 bp; see Fig. 3), the RPA technique is able to discriminate between these two allelic variants.

The selective expression of centerin in GC B cells was confirmed by *in situ* hybridization using an antisense riboprobe. As shown in Fig. 6a, the probe essentially binds to all germinal centers on a tonsil section, whereas the FM, T cell-rich, and the sub-epithelial areas fail to display centerin-specific signal, confirming that this message is up-regulated as B lymphocytes enter GC reactions. Within GC there is no apparent polarity to the staining, hence GC dark zone and light zone are indistinguishable from one another. Fig. 6b shows a consecutive tonsillar section hybridized to a sense RNA probe as

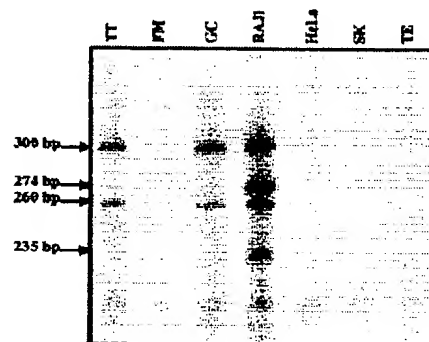


Fig. 5. RPA of total tonsil (TT), FM B cell (FM), GC B cell (GC), Burkitt's lymphoma line (RAJI), cervical carcinoma line (HeLa), neuroblastoma line (SK), and medulloblastoma line (TE) total RNA using a centerin riboprobe. Protected fragments can be seen in the TT, GC, and RAJI lanes. In the TT and GC lanes, two protected bands are present corresponding to each of the centerin splice isoforms. In the RAJI cell line, four fragments are seen: two which match the GC and TT bands, and two smaller fragments which derive from a mutant allele with a deleted valine residue at position 155.

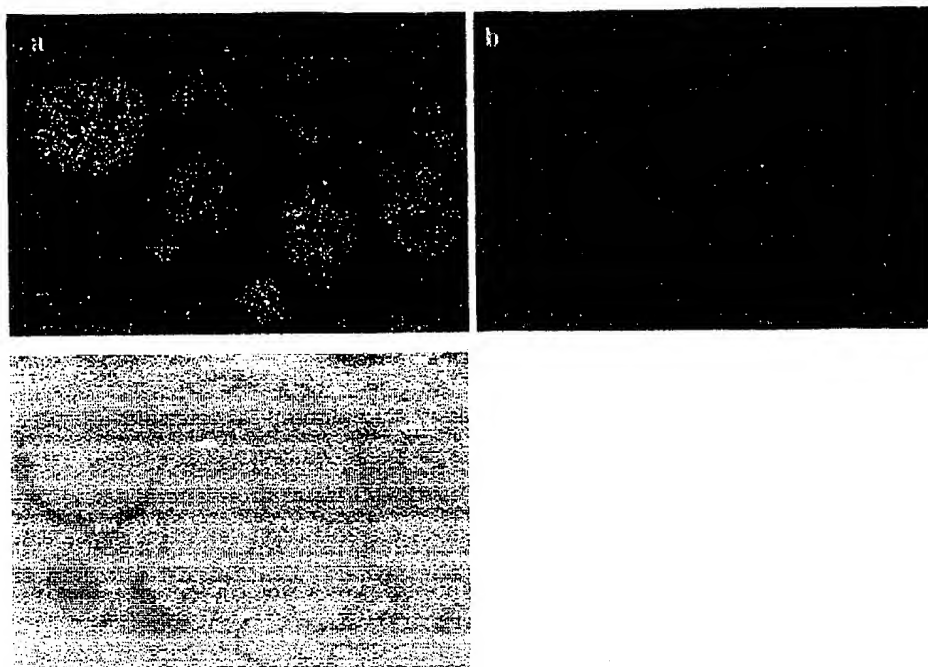


Fig. 6. *In situ* hybridization of a tonsil frozen section with centerin antisense (a) and sense (b) riboprobes. The antisense riboprobe hybridizes to the GC but not the FM areas. A serial frozen tonsillar section displaying the corresponding area weakly stained with Giemsa but without hybridization is shown in (c). Because of the higher cellular density, the FM areas are distinctively visible.

negative control, demonstrating the low background and high sensitivity of the technique. Collectively, our data indicate that the centerin message is not only transcriptionally up-regulated by GC B cells, but also highly restricted to them.

2.5 CD40L activation up-regulates centerin transcription in naive B cells

BM2 cells were purified and cultured with either SAC or sCD40L for 48 h. Cells were then harvested and used to prepare amplified cDNA representations (amplicons) as described for RDA [7, 19]. Amplicons from each cell population were then separated by electrophoresis, transferred to nylon membranes, and probed with the 612-bp RDA-isolated centerin fragment. As shown in Fig. 7, ligation of CD40 in naive B cells up-regulates the centerin message, while B cell receptor (BCR) engagement by SAC does not.

2.6 The centerin gene maps to one of the serpin clusters in chromosome 14

Using a probe corresponding to the 5' untranslated region of the centerin transcript to screen a human/hamster radiation hybrid library [25], the gene encoding this protein has been mapped to one of the major known ser-

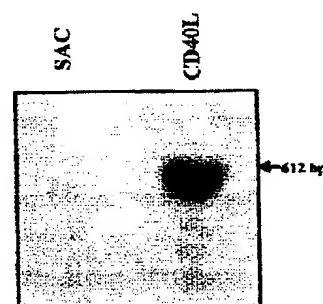


Fig. 7. Amplicon blot of SAC- and CD40L-activated naive (Bm2) B cells. Restriction-digested, amplified representations corresponding to each of the activated B cell populations were blotted and hybridized with a centerin cDNA probe. A positive band corresponding to the correct size is only visualized in the CD40L-activated B cell population.

pin clusters on chromosome 14 [26]. In fact, the centerin gene maps adjacent to protein C inhibitor (PCI), a serpin displaying homology to centerin particularly at its active site (Fig. 2) [27].

3 Discussion

Serpins are protease inhibitors involved in the regulation of multiple protease cascades. In this report, we describe the cloning of a novel member of this superfamily whose expression is restricted to GC B cells. We have thus named this novel serpin "centerin". Alignment of the deduced centerin protein sequence with several other serpin family members reveals that it possesses all of the necessary motifs to be considered an active protease inhibitor.

Serpins are structurally composed of three β -sheets surrounded by eight α -helices. The connection between A and C β -sheets constitutes the reactive site loop [21–24]. Within this site, the P1 residue plays a crucial role in determining serpin specificity. However, many serpins share identical P1 residues and yet have been shown to inhibit different target proteases. Overall, centerin's primary amino acid sequence is most homologous to the non-inhibitory carrier protein TBG, with which it shares 50% amino acid identity. However, centerin differs markedly from TBG at those sites noted above which participate in protease inhibition (Fig. 2). Centerin possesses an arginine residue at position P1, suggesting that it may inhibit trypsin-like serine proteases, which are endopeptidases that cleave C-terminal to basic arginine and lysine residues [28]. Several other serpin superfamily members possessing P1 arginine residues have been identified, including plasminogen activator inhibitor-1 and 2 (PAI-1 and PAI-2), PI-6, PI-8, protease nexin 1 (PN-1), maspin, and bomapin/PI-10, all of which differ markedly in terms of primary sequence from centerin. In addition, anti-thrombin III and PCI, which inhibit several different trypsin-like proteases of the coagulation and fibrinolytic systems, have identical arginine/serine P1-P1' reactive sites as centerin (Fig. 2) [27, 29]. Mutagenesis experiments using a variety of serpin substrates have shown, however, that the reactive site loop is not sufficient to transfer the specificity of a given serpin. Indeed, secondary binding sites and conformational constraints are responsible for the ultimate specificity of these molecules [30–32]. Our preliminary experiments suggests that recombinant centerin does not form a complex with thrombin or protein C (data not shown). Further experiments aimed at ascertaining the nature of centerin's target protease(s) are underway.

Several members of the serpin superfamily lack protease-inhibitory capability and have other physiologi-

cal roles. This last group includes, among others, TBG, angiotensinogen, ovalbumin, and pigment epithelium-derived factor (PEDF). PEDF is a neurite-promoting factor released by human retinal pigment epithelial cells which fulfils all the conformation requirements of active serpins. Its neurotrophic activity, however, is not implemented through the inhibition of a serine protease. Furthermore, the biological activity of PEDF seems to involve the N terminus of the protein rather than its serpin reactive loop [33]. PEDF has recently been described to be among the most potent natural inhibitors of angiogenesis in the eye, as it inhibits endothelial cell migration towards a wide array of angiogenic factors [34]. Whether this effect is also independent of the serpin reactive loop remains to be determined.

The centerin message is transcribed as two differentially spliced variants. The longest transcript encodes an N-terminal leader peptide, while the shortest has deleted the leader and adjacent N-terminal sequence encoding the first α -helix. Based on the three-dimensional models of α -1 anti-trypsin (AAT) and PAI-1, the first serpin α -helix is crucial to the appropriate folding of the molecule [21, 22]. It is likely, therefore, that only the longer centerin transcript encodes a functionally active serpin. Although the presence of a leader sequence in the putative 417-amino acid protein suggests an extracellular role for this molecule, there are examples of secreted serpins that play both extra- and intracellular inhibitory roles [35, 36]. Interestingly, most intracellular serpins seem to display anti-apoptotic activities. Cytokine response modifier A (CrmA), for example, is a poxvirus-encoded serpin that inhibits cytotoxic T cell-mediated apoptosis through both Ca^{2+} -dependent (inhibition of granzyme B) and Fas-dependent (inhibition of caspases 7, 10 and 3) pathways [37–42]. In addition to CrmA, there are at least three examples of serpins that control intracellular apoptosis: (1) PAI-2, a macrophage-derived serpin, protects against TNF-induced apoptosis in addition to its role in regulating extracellular urokinase, (2) PI-9, an intracellular serpin present in T lymphocytes, inhibits granzyme B and protects the cell against perforin and granzyme B-mediated killing, and (3) raPIT5a, a serpin that seems to regulate granzyme B-mediated apoptosis in the pituitary gland [35, 36, 43, 44].

The presence of centerin's aspartic acid P3' residue is noteworthy, as aspartic acid moieties are typically absent from serpin reactive center loops. However, the IL-1 β -converting enzyme family of cysteine proteases are known to cleave substrates C-terminal to aspartate amino acids, and several serpins have been shown to be inhibitory via interactions involving an acidic reactive center loop residue, including the anti-apoptotic PI-9 and CrmA [42, 43].

While undoubtedly many receptor-ligand interactions are necessary for induction of the full complement of genes necessary to initiate GC reactions, as regards B cells, two of the best-characterized examples are engagement of the surface BCR complex and CD40, by antigen and CD40L, respectively [45]. Each of these signals are known to mediate events fundamental to GC formation, and in concert can even induce many aspects of the GC B cell phenotype [46]. Consistent with its *in vivo* restricted transcription, we found that the centerin message can be up-regulated by CD40L, but not by SAC, activation of naive B cells *in vitro*. Thus, the CD40/CD40L interaction that is fundamental for GC formation [47–50], seems to provide at least one of the molecular signals capable of prompting transcription of this GC-specific, novel B cell serpin gene. Because the *in situ* hybridization pattern of centerin on tonsil sections fails to disclose any polarity within GC, it is likely that transcription of the message for this novel serpin persists during the transition from B cell centroblast to centrocyte. Finally, our PCR data confirm that the centerin message is down-regulated as these cells differentiate to memory and plasma cells.

We have detected the presence of the centerin message in the HL-60 promyelocytic leukemia line by Northern blot analysis. In the same experiments, however, we could not detect centerin transcription in bone marrow or peripheral blood RNA samples, suggesting that this gene is not expressed in either myeloid precursors, mature granulocytes or monocytes. PCR amplification of mature macrophages also failed to yield any positive. The HL-60 cell line has been described to display several chromosomal aberrations, including at least two different translocations that involve chromosome 14 where the centerin gene is located [51]. Chromosomal translocations may place this gene under abnormal transcriptional regulation, thus explaining its detection in HL-60 cells.

Several previous studies have identified links between serpin family members' functions and cellular differentiation. Centerin represents a differentially expressed gene whose transcription is induced in mature B cells that are initiating programs associated with affinity maturation and the development of B cell memory. Further investigations into the function of this protein should shed light on the biology of B cells at the time they enter the highly specialized GC environment.

4 Materials and methods

4.1 Cloning of centerin

A 612-bp fragment from the centerin message was isolated via RDA as described [6, 7, 19]. Highly pure naive and GC tonsillar B cells were used in these experiments. The initial RDA sequence was used to generate primers and extend the initial sequence 5' and 3' using rapid amplification of cDNA ends (RACE, Gibco/BRL) using Raji cell line and GC centroblast (Bm3) cDNA. The resulting products were cloned into PCR11-T-tailed vector (Invitrogen) and sequenced. Primers used in RACE reactions were: 5'-TGT GGT TCA CCA GAA CCA TGG-3'-GSP1, 5'-ATT ATG TCT ACA ACC TTC CCT-3'-GSP2, 5'-ACA TGG CTG TTG ATC CTC GCC-3'-GSP3, 5'-ATT GCC CAA GAA ATT TGC CTG-3'-GSP2b, and 5'-TGC CTG CAG CTG GAC CTC CTT-3'-GSP3b.

4.2 Northern blot

Pre-made membranes containing multiple tissue, cancer cell line, and lymphoid tissue polyA⁺ RNA (Clontech, Palo Alto, CA) were screened with [³²P]α-dCTP-labeled probes corresponding to the original centerin 612-bp RDA fragment cDNA and an actin control. The membranes were incubated with 2 × 10⁶ cpm of denatured probe per ml in ExpresshybTM hybridization solution at 68 °C for 2 h. After hybridization, membranes were washed in 2 × SSC, 0.5 % SDS followed by 0.1 × SSC, 0.1 % SDS at 50 °C for 2 h. The blots were then exposed to phosphorimager screens for 1–3 days at room temperature and subsequently visualized using phosphorimager (Molecular Dynamics, Sunnyvale, (A) and the Imagequant software package.

4.3 RNase protection assays

Total RNA from different sources was extracted using an RNA isolation kit (Qiagen). Alternatively, 1 × 10⁶–2 × 10⁶ FCM-sorted FM (IgD⁺ CD38⁺) or GC (IgD⁺ CD38⁺) cells were lysed using 100 μl of the Direct Protect lysis buffer (Ambion). The assays were conducted with either 45 μl of whole-cell lysate or 10 μg of total RNA precipitated and resuspended in 45 μl of lysis buffer. An antisense RNA probe corresponding to a 612-bp fragment of centerin was prepared using *in vitro* transcription, by incubation at 37 °C for 1 h in a 20-μl reaction containing 1 μg template DNA, 0.5 mM GTP, 0.5 mM ATP, 0.5 mM UTP, and 62.5 μCi [³²P]dCTP, 2 U of either SP6 or T7 RNA polymerase, and 1 × transcription buffer (Maxi-script/Ambion). The template was subsequently removed by the addition of 2 U DNase I for 15 min. Unincorporated nucleotides were removed using a micro-Bio-spin column (Bio-Rad). Typically, 10⁵ of the riboprobe (specific activity of 10⁸ cpm/μg) was added to the cell lysate or RNA, and the mixture incubated overnight at 37 °C. The mixture was then

treated with RNase A/T1 cocktail (Ambion Direct Protect Kit) for 30 min at 37 °C, followed by proteinase K treatment. The protected material was precipitated with isopropanol, heated to 90 °C, and loaded onto 5 % polyacrylamide gels at 600 V for 2–3 h. The gels were then transferred to paper and visualized using a Molecular Dynamics phosphorimager. GAPDH was used as an RNA loading control.

4.4 Analyses of DNA sequences

Sequences were edited and analyzed using the DNASTar software package (DNASTar Inc., Madison, WI). Cloned products were searched against the NR (DNA and protein) and EST (DNA) databases administered by the National Center of Biotechnology Information (NCBI) using the Basic and Advanced Local Alignment Search Tool (BLAST) algorithm [20].

4.5 PCR amplification of centerin

The PCR primers 5'-TTT TCT TTA AAG CCA AGT GGG AGA-3' and 5'-AAA ATC TGG GGA TGA ACA CCT CTA-3' were used to amplify a fragment of the centerin gene that spans an intron, thus preventing DNA co-amplification. Total RNA was extracted from (i) human macrophages obtained after culturing the adherent PBMC fraction for 6 days in the presence of GM-CSF, (ii) the human myeloma cell lines U266 and RPMI 8221, (iii) enriched CD20⁺ peripheral blood B cells negatively selected after removal of CD3⁺, CD14⁺, and CD56⁺ cells from PBMC (Miltenyi Biotec), and (iv) enriched GC B cells obtained after positive selection of CD20⁺ CD38⁺ tonsillar B cells using magnetic beads (Dynabeads, Dynal).

4.6 *In situ* hybridization

The N-terminal 541 bp of centerin were cloned into the pCRII vector (Invitrogen), which contains both T7 and SP6 RNA polymerase promoters. Centerin was amplified from this vector by PCR using the M13 forward and M13 reverse primers, which flank the RNA polymerase promoter sites. The centerin PCR product was then precipitated and subsequently used as the template for *in vitro* transcription. Probe was prepared by incubation at 37 °C for 1 h in a 20- μ l reaction containing 1 μ g of template DNA, 0.5 mM CTP, 0.5 mM GTP, 0.5 mM ATP, 0.005 mM UTP, 62.5 μ Ci [³²S]UTP, 2 U of either SP6 or T7 RNA polymerase, and 1x transcription buffer (Maxiscript/Ambion). The template was subsequently removed by the addition of 2 U DNase I for 15 min. Unincorporated nucleotides were removed using a micro-Bio-spin column (Bio-Rad).

The tonsils were fixed in 4 % paraformaldehyde-PBS overnight at 4 °C. The tissue was placed in 70 % ethanol, dehydrated through graded ethanol solutions, cleared with xylene and infused with paraffin. Contiguous sections were

probed with sense and antisense transcripts or stained, as described previously [52].

4.7 Generation of amplicon blots from SAC- and CD40L-activated naive B cells

The cell substrates used in RDA were kindly provided by Dr. Francine Briere (Schering-Plough, Dardilly, France). Naive (Bm2) cells were purified from human tonsils based on the expression of IgD and CD23. Purified cells were resuspended at 5×10^5 cells/ml in RPMI 1640 medium with 10 % FCS. A total of 100–200 μ l of the suspension was dispensed in microwells. SAC particles (0.001–0.005 %, v/v) or sCD40L were added to the cultures and the cells were harvested after 48 h. Amplicons were constructed as described [6, 7] after restriction digestion and PCR amplification (18 cycles) of the cDNA corresponding to each of the activated B cell populations. Amplicon blots were hybridized with a ³²P-labeled double-stranded cDNA probe generated by PCR radioactive labeling (Life Technologies), followed by stringent washing, and developed using a PhosphorImager and the Imagequant software package (Molecular Dynamics).

4.8 Chromosomal localization of the centerin gene

Centerin-specific primers (5'-GCCAATGCCCCAGTGCA-TAC-3' and 5'-AGGTCAGTCTTTGCTGGGAACAG-3') were used in a radiation hybrid (RH) screening via PCR on the Genebridge 4 panel (Research Genetics, Huntsville, AL). Briefly, 25 ng DNA from each of the 93 cell lines were used as PCR template for the primer combination to be screened. Using standard conditions, PCR products were assayed on 3 % agarose gels and the results scored for the presence (1) or absence (0) of a product of the expected size. The data was submitted to the Whitehead/MIT RH server where it was tested against recently published framework maps [25].

Acknowledgements: The authors are grateful to Dr. Francine Briere for kindly providing the SAC- and CD40L-activated BM2 cells and to Odette de Bouteiller, Steven Scholl, and Jeanette LeGros for excellent technical assistance. This work was supported in part by National Institutes of Health grants nos. AI-42862 and AI-12127 and grants from the Baylor Health Care System Foundation.

References

- 1 Rajewsky, K., Clonal selection and learning in the antibody system. *Nature* 1996. 381: 751–758.
- 2 MacLennan, I. C., Germinal centers. *Annu. Rev. Immunol.* 1994. 12: 117–139.
- 3 Kelsoe, G., *In situ* studies of the germinal center action. *Adv. Immunol.* 1995. 60: 267–288.

- 4 Liu, Y.-J. and Banchereau, J., The paths and molecular controls of peripheral B cell development. *Immunologist* 1996. 4: 55–66.
- 5 Liu, Y.-J. and Arpin, C., Germinal center development. *Immunol. Rev.* 1997. 156: 111–126.
- 6 Frazer, J. K., LeGros, J., de Bouteiller, O., Liu, Y.-J., Banchereau, J., Pascual, V. and Capra, J. D., Identification and cloning of genes expressed by human tonsillar B lymphocyte subsets. *Ann. N. Y. Acad. Sci.* 1997. 815: 316–318.
- 7 Frazer, J. K., Pascual, V. and Capra, J. D., RDA of lymphocyte subsets. *J. Immunol. Methods* 1997. 207: 1–12.
- 8 Hunt, L. T. and Dayhoff, M. O., A surprising new protein superfamily containing ovalbumin, antithrombin-III, and alpha 1-proteinase inhibitor. *Biochem. Biophys. Res. Commun.* 1980. 95: 864–871.
- 9 Potempa, J., Korzus, E. and Travis, J., The serpin superfamily of proteinase inhibitors: structure, function, and regulation. *J. Biol. Chem.* 1994. 269: 15957–15960.
- 10 Stone, S. R., Whisstock, J. C., Bottomley, S. P. and Hopkins, P. C., Serpins. A mechanistic class of their own. *Adv. Exp. Med. Biol.* 1997. 425: 5–15.
- 11 Stein, P. E., Leslie, A. G., Finch, J. T., Turnell, W. G., McLaughlin, P. J. and Carrell, R. W., Crystal structure of ovalbumin as a model for the reactive centre of serpins. *Nature* 1990. 347: 99–102.
- 12 Patston, P. A., Gettins, P., Beechem, J. and Schapira, M., Mechanism of serpin action: evidence that C1 inhibitor functions as a suicide substrate. *Biochemistry* 1991. 30: 8876–8882.
- 13 Matheson, N. R., van Halbeek, H. and Travis, J., Evidence for a tetrahedral intermediate complex during serpin-proteinase interactions. *J. Biol. Chem.* 1991. 266: 13489–13491.
- 14 Schulze, A. J., Huber, R., Bode, W. and Engh, R. A., Structural aspects of serpin inhibition. *FEBS Lett* 1994. 344: 117–124.
- 15 Wright, H. T., The structural puzzle of how serpin serine proteinase inhibitors work. *Bioessays* 1996. 18: 453–464.
- 16 Lawrence, D. A., The serpin-proteinase complex revealed. *Nat. Struct. Biol.* 1997. 4: 339–341.
- 17 Flink, I. L., Bailey, T. J., Gustafson, T. A., Markham, B. E. and Morkin, E., Complete amino acid sequence of human thyroxine-binding globulin deduced from cloned DNA: close homology to the serine antiproteases. *Proc. Natl. Acad. Sci. USA* 1986. 83: 7708–7712.
- 18 Hayashi, Y., Mori, Y., Janssen, O. E., Sunthornthepvarakul, T., Weiss, R. E., Takeda, K., Weinberg, M., Seo, H., Bell, G. I. and Refetoff, S., Human thyroxine-binding globulin gene: complete sequence and transcriptional regulation. *Mol. Endocrinol.* 1993. 7: 1049–1060.
- 19 Hubank, M. and Schatz, D. G., Identifying differences in mRNA expression by representational difference analysis of cDNA. *Nucleic Acids Res.* 1994. 22: 5640–5648.
- 20 Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997. 25: 3389–3402.
- 21 Huber, R. and Carrell, R. W., Implications of the three-dimensional structure of α -antitrypsin for structure and function of serpins. *Biochemistry* 1989. 28: 8951–8966.
- 22 Mottonen, J., Strand, A., Symersky, J., Sweet, R. M., Danley, D. E., Geoghegan, K. F., Gerard, R. D. and Goldsmith, E. J., Structural basis of latency in plasminogen activator inhibitor-1. *Nature* 1992. 355: 270–273.
- 23 Wei, A., Rubin, H., Cooperman, B. S. and Christianson, D. W., Crystal structure of an uncleaved serpin reveals the conformation of an inhibitory reactive loop. *Nat. Struct. Biol.* 1994. 1: 251–258.
- 24 Ryu, S. E., Choi, H. J., Kwon, K. S., Lee, K. N. and Yu, M. H., The native strains in the hydrophobic core and flexible reactive loop of a serine protease inhibitor: crystal structure of an uncleaved α 1-antitrypsin at 2.7 Å. *Structure* 1996. 4: 1181–1192.
- 25 Hudson, T. J., Stein, L. D., Gerety, S. S., et al., An STS-based map of the human genome. *Science* 1995. 270: 1945–1954.
- 26 Billingsley, G. D., Walter, M. A., Hammond, G. L. and Cox, D. W., Physical mapping of four serpin genes: α 1-antitrypsin, α 1-antichymotrypsin, corticosteroid-binding globulin, and protein C inhibitor, within a 280-kb region on chromosome 14q32.1. *Am. J. Hum. Genet.* 1993. 52: 343–353.
- 27 Suzuki, K., Deyashiki, Y., Nishiooka, J., Kurachi, K., Akira, M., Yamamoto, S. and Hashimoto, S., Characterization of a cDNA for human protein C inhibitor. A new member of the plasma serine protease inhibitor superfamily. *J. Biol. Chem.* 1987. 262: 611–616.
- 28 Olson, S. T., Bock, P. E., Kvaasman, J., Shore, J. D., Lawrence, D. A., Ginsburg, D. and Björk, I., Role of the catalytic serine in the interactions of serine proteinases with protein inhibitors of the serpin family. *J. Biol. Chem.* 1995. 270: 30007–30017.
- 29 Chandra, T., Stackhouse, R., Kidd, V. J. and Woo, S. L., Isolation and sequence characterization of a cDNA clone of human antithrombin III. *Proc. Natl. Acad. Sci. USA* 1983. 80: 1845–1848.
- 30 Djie, M. Z., Stone, S. R. and Le Bonniec, B. F., Intrinsic specificity of the reactive site loop of α 1-antitrypsin, α 1-antichymotrypsin, antithrombin III, and protease nexin I. *J. Biol. Chem.* 1997. 272: 16268–16273.
- 31 Bottomley, S. P. and Stone, S. R., Protein engineering of chimeric serpins: an investigation into effects of the serpin scaffold and reactive centre loop length. *Protein Eng.* 1998. 11: 1243–1247.
- 32 Djie, M. Z., Le Bonniec, B. F., Hopkins, P. C., Hipler, K. and Stone, S. R., Role of the P2 residue in determining the specificity of serpins. *Biochemistry* 1996. 35: 11461–11469.
- 33 Becerra, S. P., Sagasti, A., Spinella, P. and Notario, V., Pigment epithelium-derived factor behaves like a noninhibitory serpin. Neurotrophic activity does not require the serpin reactive loop. *J. Biol. Chem.* 1995. 270: 25992–25999.
- 34 Dawson, D. W., Volpert, O. V., Gillis, P., Crawford, S. E., Xu, H., Benedict, W. and Bouck, N. P., Pigment epithelium-derived factor: a potent inhibitor of angiogenesis. *Science* 1999. 285: 245–248.
- 35 Dickinson, J. L., Bates, E. J., Ferrante, A. and Antalis, T. M., Plasminogen activator inhibitor type 2 inhibits tumor necrosis factor- α -induced apoptosis. Evidence for an alternate biological function. *J. Biol. Chem.* 1995. 270: 27894–27904.
- 36 Antalis, T. M., La Linn, M., Donnan, K., Mateo, L., Gardner, J., Dickinson, J. L., Buttigieg, K. and Suhrbier, A., The serine proteinase inhibitor (serpin) plasminogen activation inhibitor type 2 protects against viral cytopathic effects by constitutive interferon alpha/beta priming. *J. Exp. Med.* 1998. 187: 1799–1811.
- 37 Komiya, T., Ray, C. A., Pickup, D. J., Howard, A. D., Thornberry, N. A., Peterson, E. P. and Salvesen, G., Inhibition of interleukin-1 β converting enzyme by the cowpox virus serpin CrmA. An example of cross-class inhibition. *J. Biol. Chem.* 1994. 269: 19331–19337.
- 38 Quan, L. T., Caputo, A., Bleackley, R. C., Pickup, D. J. and Salvesen, G. S., Granzyme B is inhibited by the cowpox virus serpin cytokine response modifier A. *J. Biol. Chem.* 1995. 270: 10377–10379.

- 39 Tewari, M., Telford, W. G., Miller, R. A. and Dixit, V. M., CmA, a poxvirus-encoded serpin, inhibits cytotoxic T-lymphocyte-mediated apoptosis. *J. Biol. Chem.* 1995. **270**: 22705-22708.
- 40 Tewari, M. and Dixit, V. M., Fas- and tumor necrosis factor-induced apoptosis is inhibited by the poxvirus cmA gene products. *J. Biol. Chem.* 1995. **270**: 3255-3260.
- 41 Takahashi, A., Muey, P. Y., Martins, L. M., Poirier, G. G., Moyer, R. W. and Earnshaw, W. C., CmA/SPI-2 inhibition of an endogenous ICE-related protease responsible for lamin A cleavage and apoptotic nuclear fragmentation. *J. Biol. Chem.* 1996. **271**: 32487-32490.
- 42 Zhou, Q., Snipas, S., Orth, K., Muzio, M., Dixit, V. M. and Salvesen, G. S., Target protease specificity of the viral serpin CmA. Analysis of five caspases. *J. Biol. Chem.* 1997. **272**: 7797-7800.
- 43 Sun, J., Bird, C. H., Sutton, V., McDonald, L., Coughlin, P. B., De Jong, T. A., Trapani, J. A. and Bird, P. I., A cytosolic granzyme B inhibitor related to the viral apoptotic regulator cytokine response modifier A is present in cytotoxic lymphocytes. *J. Biol. Chem.* 1996. **271**: 27802-27809.
- 44 Hill, R. M., Morressey, K. S., Coates, L. C., Mezey, E., Fell, B., Bratt, T., Trapani, J. A. and Birch, N. P., A new intracellular serine protease inhibitor expressed in the rat pituitary gland complexes with granzyme B. *FEBS Lett.* 1998. **440**: 361-364.
- 45 Briere, F. and Banchereau, J., Polyclonal activation of human mature B lymphocytes. In Herzenberg, L. A., Weir, D. M., Herzenberg, L. A. and Blackwell, B. (Eds.) *Handbook of Experimental Immunology*, 5th edn, Vol. IV, section 33. Blackwell, Oxford 1997, p. 224.1.
- 46 Galibert, L., Burdin, L., Banchereau, J. and Rousset, F., CD40 and BCR dual triggering of resting B lymphocytes turns on a partial GC phenotype. *J. Exp. Med.* 1996. **183**: 77-85.
- 47 Allen, R. S., Armitage, R. J., Conley, M. E., et al., CD40 ligand gene defects responsible for X-linked hyper-IgM syndrome. *Science* 1993. **259**: 990-993.
- 48 DiSanto, J. P., Bonnefoy, J. Y., Gauchat, J. F., Fischer, A. and de Saint Basile, G., CD40 ligand mutations in X-linked immunodeficiency with hyper-IgM. *Nature* 1993. **361**: 541-543.
- 49 Aruffo, A., Farrington, M., Hollenbaugh, D., et al., The CD40 ligand, gp39, is defective in activated T cells from patients with X-linked hyper-IgM syndrome. *Cell* 1993. **72**: 291-300.
- 50 Korthauer, U., Graf, D., Mages, H. W., Briere, F., Padayachee, M., Malcolm, S., Ugazio, A. G., Notarangelo, L. D., Levinsky, R. J. and Kroczeck, R. A., Defective expression of T-cell CD40 ligand causes X-linked immunodeficiency with hyper-IgM. *Nature* 1993. **361**: 539-541.
- 51 Liang, J. C., Ning, Y., Wang, R. Y., Padilla-Nash, H. M., Schrock, E., Soenksen, D., Nagarajan, L. and Ried, T., Spectral karyotypic study of the HL-60 cell line: detection of complex rearrangements involving chromosomes 5, 7, and 16 and delineation of critical region deletion on 5q31.1. *Cancer Genet. Cytogenet.* 1999. **113**: 105-109.
- 52 Benjamin, I. J., Shelton, J., Garry, D. J. and Richardson, J. A., Temporospatial expression of the small HSP/aB-crystallin in cardiac and skeletal muscle during development. *Dev. Dyn.* 1997. **208**: 75-84.

Correspondence: Virginia Pascual, Baylor Institute for Immunology Research, 3434 Live Oak, Dallas, TX 75204, USA

Fax: +1-214-820-4813

e-mail: virginip@bayloridallas.edu